

# **Molecular Mapping of Genes Involved in Carbohydrate Metabolism and Fluorescent AFLP-based Tagging of QTL in Tetraploid Potato**

---



**Inaugural-Dissertation  
zur  
Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität zu Köln**

**Vorgelegt von  
CHEN Xinwei  
Aus Peking, China**

**Köln 2000**



Die vorliegende Arbeit wurde am Max-Planck-Institut für Züchtungsforschung in Köln- Vogelsang, in der Abteilung Pflanzenzüchtung und Ertragsphysiologie (Prof. F. Salamini) in der Arbeitsgruppe von Prov. Doz. Dr. Christiane Gebhardt.

Berichterstatter: PD Dr. Christiane Gebhardt  
Prof. Dr. Diethard Tautz  
Vorsitzender: Prof. Dr. Ulf-Ingo Flügge

Tag der mündlichen Prüfung: 08/11/2000



*for my father*



## Abbreviations

<b>AFLP</b>	amplified fragment length polymorphism
<b>Amp</b>	ampicillin
<b>ANOVA</b>	analysis of variance
<b>APS</b>	ammonium persulphate
<b>bp</b>	base pairs
<b>BSA</b>	bovine serum albumin / bulked segregant analysis
<b>cDNA</b>	complementary deoxyribonucleic acid
<b>CAPS</b>	cleaved amplified polymorphic sequence
<b>cM</b>	centiMorgans
<b>dATP</b>	deoxyadenosinetriphosphate
<b>dCTP</b>	deoxycytosinetriphosphate
<b>dGTP</b>	deoxyguanosinetriphosphate
<b>dTTP</b>	deoxythymidinetriphosphate
<b>DH</b>	doubled haploid
<b>DTT</b>	dithiothreitol
<b><i>E. coli</i></b>	<i>Escherichia coli</i>
<b>EDTA</b>	ethylenediamine tetraacetic acid
<b>EST</b>	expressed sequence tag
<b>EtBr</b>	ethidium bromide
<b>EtOH</b>	ethanol
<b>fAFLP</b>	fluorescence-labelled AFLP
<b>FE</b>	foliage earliness
<b>GLM</b>	general linear models
<b>h</b>	hour(s)
<b>IPTG</b>	isopropylthio- $\beta$ -o-galactopyranoside
<b>kb</b>	kilo base
<b>LOD</b>	logarithm of odds
<b>M</b>	Molar
<b>MAS</b>	marker-assisted selection
<b>min</b>	minute(s)
<b>PCR</b>	polymerase chain reaction
<b>QTL</b>	quantitative trait loci
<b>rAFLP</b>	radiolabelled AFLP
<b>RAPD</b>	randomly amplified polymorphic DNA
<b>RFLP</b>	restriction fragment length polymorphism
<b>RIL</b>	recombinant inbred line
<b>RSC</b>	reducing sugar content
<b>s</b>	second(s)
<b>SCAR</b>	sequence characterised amplified region
<b>SDS</b>	sodium dodecyl sulfate
<b>SNP</b>	single nucleotide polymorphism
<b>STS</b>	sequence tagged site
<b>SSR</b>	simple sequence repeat
<b>SSCP</b>	single-strand conformational polymorphism
<b>SY</b>	starch yield
<b>TSC</b>	tuber starch content
<b>TY</b>	tuber yield
<b>X- gal</b>	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside

# CONTENTS

---

## 1. INTRODUCTION

1.1 Taxonomy and importance of the potato .....	1
1.2 Evolution of cultivated potato.....	1
1.3 History of the cultivated potato .....	2
1.4 Genetic features of the potato .....	2
1.5 Mapping in plant.....	4
1.5.1 Mapping populations.....	4
1.5.2 Molecular marker technologies for Genetic mapping.....	6
1.5.3 Application of molecular genetic mapping .....	11
1.6 Function mapping.....	12
1.6.1 Metabolic pathways relevant to carbohydrate content in potato tubers .....	13
1.6.1.1 The Calvin cycle.....	14
1.6.1.2 Respiratory pathways .....	14
1.6.1.3 Sucrose metabolism.....	17
1.6.1.4 Starch metabolism .....	17
1.6.1.5 Major transporters for carbon exchange between the cytosol and the plastid .....	18
1.7 QTL analysis.....	19
1.7.1 Approaches to QTL analysis .....	19
1.7.1.1 QTL analysis through a molecular marker approach .....	19
1.7.1.2 QTL analysis through a candidate gene approach .....	20
1.7.2 QTL analysis in potato.....	21
1.8 Agronomic traits under investigation .....	21
1.9 Outline of the thesis.....	22

## 2. MATERIALS AND METHODS

2.1 Chemicals, enzymes and oligonucleotides .....	24
---	----

---

---

2.2 Plant material.....	24
2.3 Bacterial strain and cloning vector .....	25
2.4 Media .....	25
2.5 Buffers and solutions .....	26
2.6 DNA extraction.....	27
2.6.1 Plant genomic DNA extraction .....	27
2.6.2 DNA extraction from agarose gel .....	27
2.6.3 DNA extraction from polyacrylamide gel.....	27
2.6.4 Plasmid DNA extraction .....	28
2.6.5 Concentration measurement of DNA .....	28
2.7 Cloning PCR and AFLP fragments.....	28
2.7.1 Ligation of fragments in pGEM <sup>o</sup> -T.....	28
2.7.2 Transformation of <i>E. coli</i> .....	29
2.8 DNA sequencing and sequence analysis .....	30
2.9 Restriction fragment length polymorphism (RFLP).....	30
2.9.1 Genomic DNA restriction and purification .....	30
2.9.2 Fragment separation by electrophoresis .....	30
2.9.3 Electro-blotting of DNA fragments .....	31
2.9.4 Radioactive labelling of DNA probes.....	31
2.9.5 Hybridisation.....	32
2.10 Sequence-tagged sites (STS).....	32
2.10.1 Exploration of functional gene sequences in databases .....	32
2.10.2 Nomenclature of the functional genes.....	33
2.10.3 Polymerase chain reaction (PCR).....	33
2.10.4 Identification of sequence characterized amplified region (SCAR)....	
2.10.5 Identification of cleaved amplified polymorphic sequence (CAPS)....	
2.11 Generation of segregation data and linkage analysis.....	35
2.12 Detection of single nucleotide polymorphisms (SNPs).....	35
2.13 Amplified fragment length polymorphism (AFLP).....	35
2.13.1 Radiolabelled amplified fragment length polymorphism (rAFLP) analysis .....	35
2.13.2 Fluorescence-labelled amplified fragment length polymorphism	

---

(fAFLP) analysis .....	39
2.14 Evaluation of phenotypic traits .....	44
2.15 QTL analysis .....	45

### 3. RESULTS

3.1 Mapping functional genes involved in carbohydrate metabolism and transport .....	47
3.1.1 PCR amplification of functional genes through an STS approach ..	47
3.1.2 Polymorphism identification.....	49
3.1.3 Segregation analysis and function map construction.....	56
3.1.4 Homology of PCR products to the reference sequences.....	58
3.1.5 An STS approach detected similar map positions as revealed by RFLP .....	60
3.1.6 Overview of the potato function map .....	61
3.2 Identification of single nucleotide polymorphisms (SNPs) in <i>Stp23</i> .....	61
3.3 Molecular tagging of QTL in tetraploid potato by fAFLP markers.....	66
3.3.1 Genotyping of Z3 family by an automated fAFLP system.....	66
3.3.1.1 Screen primer combinations for fAFLP .....	66
3.3.1.2 fAFLP genotyping and data collection .....	67
3.3.2 Analysis of the phenotypic traits and phenotypic data .....	70
3.3.3 Identification of QTL associated with fAFLP markers.....	72
3.3.3.1 Association of fAFLP fragments with TSC. ....	74
3.3.3.2 Association of fAFLP fragments with RSC and FE .....	75
3.3.3.3 Association of fAFLP fragments with multiple traits .....	79
3.3.3.4 Association analysis on 50 selected genotypes in trial 1999 .....	80
3.3.4 Analysis of the segregation ratios in population Z3 .....	80
3.3.5 Assignment of map positions to QTL-linked fAFLP markers.....	82
3.3.6 QTL analysis through candidate gene strategy .....	82
3.3.7 Cloning of fAFLP fragments .....	85

### 4. DISCUSSION

4.1 Mapping functional genes involved in carbohydrate metabolism and	
--	--

transport .....	88
4.1.1 STS as an approach for functional gene mapping .....	88
4.1.2 Validity of STS approach for mapping functional genes .....	89
4.2 The potato function map.....	90
4.2.1 PCR-based markers for genetic analysis and marker assisted selection.....	90
4.2.2 Molecular function maps for comparative studies of plant genomes...	
4.2.3 Candidate genes for tuber starch content.....	92
4.3 Single nucleotide polymorphism (SNP).....	95
4.4 QTL analysis.....	98
4.4.1 Strategy of QTL tagging.....	98
4.4.2 Magnitude of fAFLP-tagged QTL effects .....	101
4.4.3 Assignment of map positions of fAFLPs and segregation analysis	102
4.4.4 Genetic basis of correlations between phenotypic traits.....	104
4.4.5 Cloning and conversion of fAFLP markers and comparison of the two genotyping systems fAFLP and rAFLP .....	106
4.4.6 Function map-based QTL analysis .....	108
<b>5. SUMMARY .....</b>	<b>109</b>
5.1 Function mapping as an approach to the identification of genes controlling tuber starch content in potato .....	110
5.2 fAFLP tagging of genetic factors controlling TSC, RSC, and FE in tetraploid potato.....	111
<b>6. REFERENCES .....</b>	<b>115</b>
<b>7. APPENDIX .....</b>	<b>130</b>

## LIST OF FIGURES

1-1 Evolutionary relationships of cultivated potatoes and their ploidy levels .....	2
---	---



---

1-2 Common strategies for the construction of mapping populations .....	5
1-3 Schematic representation of AFLP .....	8
3-1 Multiple alignment of partial cDNA sequences of the <i>Pgk</i> gene from different species .....	49
3-2 CAPS marker for the $\alpha$ -glucan phosphorylase gene <i>Stp23</i> .....	50
3-3 Molecular function map of potato .....	57
3-4 Comparison of the nucleotide sequence between potato PCR product of <i>Pgk</i> gene and one of the reference cDNA sequences of the gene .....	59
3-5 The alignment of partial sequences of <i>Stp23</i> showing the features of SNPs between different alleles of different potato lines .....	64
3-6 Partial electropherogram of nucleotide sequence data from a PCR product of gene <i>Stp23</i> showing the different types of SNP .....	65
3-7 Electropherogram and gel view of fAFLP samples run on the ABI PRISM <sup>®</sup> 377 DNA sequencer .....	68
3-8 GeneScan <sup>®</sup> -generated fAFLP electropherogram .....	69
3-9 Frequency distribution of phenotypic scores for traits TSC, RSC and FE .....	71
3-10 Map positions of fAFLP markers associated with phenotypic traits .....	83
3-11 The combined consensus sequence from the 7 sequences of the cloned fAFLP fragment B3-215 .....	86
3-12 Extended nucleotide sequence of the cloned fAFLP fragment B3-215 .....	87

## LIST OF TABLES

1-1 Allele configurations in diploid and tetraploid potato and their progeny .....	3
1-2 Phenotypic segregation ratios of a dominant allele <i>A</i> for the crosses of	

---

---

tetraploid genotypes.....	3
1-3 Major known enzymes involved in carbohydrate metabolism and transport .....	15
2-1 Adaptor and primer sequences and fluorescent labelling dye used for fAFLP analysis .....	39
3-1 Genes placed onto the potato molecular function map .....	51
3-2 Sequence similarity between the PCR products from potato and the corresponding reference sequences.....	58
3-3 CAPS markers of seven genes localised to similar map positions as mapped by RFLP.....	60
3-4 Number of polymorphic versus total AFLP fragments generated by primer sets <i>EcoRI/TaqI</i> and <i>HindIII/TaqI</i> .....	67
3-5 Number of polymorphic versus total AFLP fragments generated by <i>EcoRI/MseI</i> primer sets .....	67
3-6 Number of fAFLP markers generated by seven primer combinations in population Z3 and their origins .....	70
3-7 Heritability and correlation coefficient of the phenotypic traits evaluated.....	71
3-7 fAFLP markers showing associations with QTL for TSC as detected by the t-test and regression analysis .....	73
3-8 fAFLP markers showing associations with QTL for RSC and FE as detected by the Mann-Whitney U-test.....	75
3-10 fAFLP markers showing associations with multiple phenotypic traits .....	77
3-11 fAFLP markers associated with phenotypic traits as detected by the U-test based on 50 selected genotypes in trial 1999 .....	81
3-12 Functional gene markers showing associations with multiple phenotypic traits .....	84

---

# 1. INTRODUCTION

---

## 1.1 Taxonomy and importance of the potato

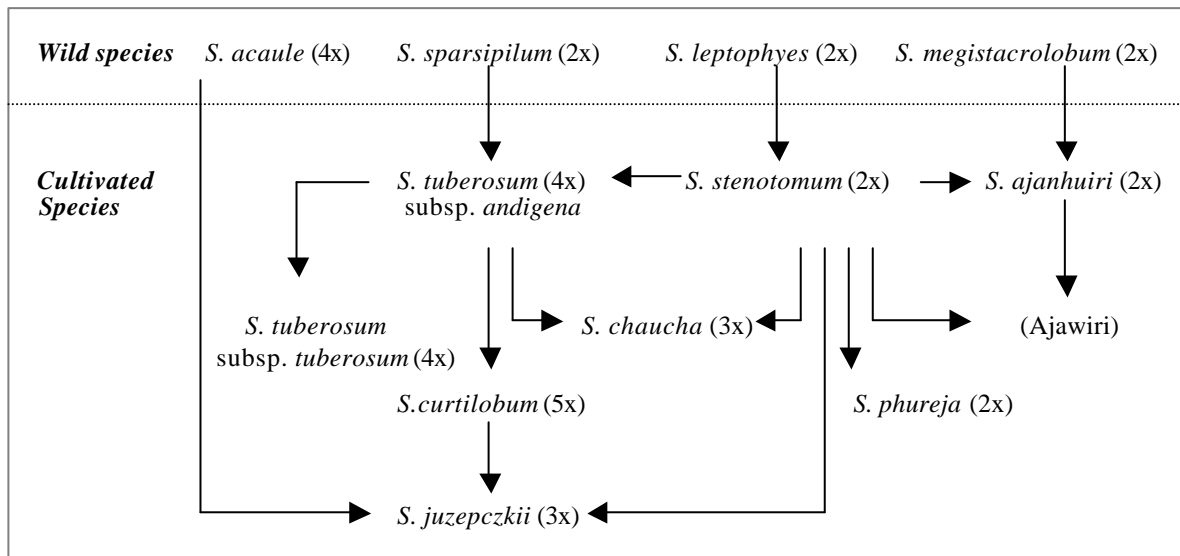
Cultivated potato (*Solanum tuberosum* ssp. *tuberosum*) belongs to the large and diverse genus *Solanum* of the family *Solanaceae*. *Solanaceae* is one of the largest and economically most important families of angiosperms. Besides potato, other economically important species of the family are tomato (*Lycopersicon esculentum*), egg-plant or aubergine (*Solanum melongena*), the garden peppers (*Capsicum annuum*) tobacco (*Nicotiana tabacum*) and pepino (*Solanum muricatum*). A number of ornamentals, petunia (*Petunia hybrida*) for example, also belong to the *Solanaceae*.

As a food crop, potato is one of the most important crops worldwide ranking 4<sup>th</sup> in terms of total world production after wheat, maize and rice. Potato tubers are valued for their high starch content. Starch can account for 80 % of the dry-weight of mature potato tubers (Kruger 1997), thus, it is an especially rich source of energy. Compared to cereals, potato yields on average more food energy and protein per unit of land (Horton 1988). Potato tubers also contain nutritive compounds such as high-quality protein (2 % of fresh weight) and substantial amounts of essential vitamins, minerals and trace elements. The production of protein per unit area of crop is second only to soybean (Tarn *et al.* 1992). The lysine content of potato can complement cereal-based diets that are deficient in this amino acid (Ortiz 1998). The importance of potato has been demonstrated by the devastating effect of the Irish famine in the 1840s. The continuing importance is reflected by the increasing growth rate of potato production, which has exceeded that of most other food crops in recent years in developing countries, particularly in Asia (Maldonado *et al.* 1998).

## 1.2 Evolution of cultivated potato

Potato is one of as many as 235 tuber-bearing *Solanum* species recognized to date, of which 228 are wild and 7 are cultivated (Hawkes 1990). Four wild species including *S. acaule*, *S. sparsipilum*, *S. leptophyes*, and *S. megistacrolobum* were believed to be involved in the evolution of the 7 cultivated species including *S. tuberosum* ssp. *andigena*, *S. tuberosum* ssp. *tuberosum*, *S. curtilobum*, *S. juzepczukii*, *S. stenotomum*, *S. chaucha*, *S. ajanhuiri*, and *S. phureja*. They occur in a range of chromosome numbers from 24 (diploid), 36 (triploid), 48

(tetraploid) to 60 (pentaploid). The evolutionary relationships are schematically presented in Figure 1-1.



**Figure 1-1** Evolutionary relationships of cultivated potatoes and their ploidy levels (adapted from Hawkes, 1990)

### 1.3 History of the cultivated potato

Potato is one of the most important world crops, yet until the 16<sup>th</sup> century it was restricted to South America and remained unknown to the people of Europe, Asia, Africa and North America. It is generally believed that potato was introduced to Europe as early as 1565 (Thornton, 1980) from Northern Colombia and further casually introduced in the 17<sup>th</sup> and 18<sup>th</sup> centuries (Bradshaw and Mackay 1994). Potato was then introduced to China, India and Japan and part of Africa in the 17<sup>th</sup> century. The first European potatoes were short-day adapted tetraploid potatoes (*Solanum tuberosum* ssp. *andigena*) and restricted to the South of Spain and Italy (Hawkes, 1978). After a few centuries of natural and unconscious selection in Europe, plant could produce tubers under the long-day photoperiodic condition and became a popular crop in other parts of Europe. The cultivated potato in Europe today has the botanical name *Solanum tuberosum* ssp. *tuberosum*.

### 1.4 Genetic features of the potato

The cultivated potato is represented by both tetraploid and diploid forms with four and two homologous sets of 12 chromosomes, respectively. Most diploid potatoes (*S. tuberosum* ssp. *tuberosum*) exhibit self-incompatibility of the gametophytic type, in which the pollen self-incompatibility is determined by the haploid genotype of the pollen. This self-incompatibility is controlled by the single, multi-allelic locus S (De Nettancourt 1977, Kaufmann *et al.* 1991)

## Introduction

which was mapped on chromosome I (Gebhardt *et al.* 1991). Tetraploid potato and other polyploid species are generally self-compatible because of the action of more than one S allele in the pollen grain (Crane *et al.* 1929). They suffer, however, inbreeding depression after a few generations of selfing. Consequently, potatoes are highly outcrossing and highly heterozygous with multiple alleles at virtually every locus. This implies that random pairing

**Table 1-1 Allele configurations in diploid and tetraploid potato and their progeny**

**A**

? \ ?	<b><i>A1*</i></b>	<b><i>A2</i></b>
<b><i>A3</i></b>	A1A3	A2A3
<b><i>A4</i></b>	A1A4	A2A4

\*A1, A2, A3, ...A8 are different alleles at one locus. For two given parental lines the number of segregating alleles could be as high as 4 and 8 for diploid and tetraploid plant, respectively. The possible genotypes of gametes (in bold italics) and zygotes are indicated as in **A** for diploid and **B** for tetraploid plants without considering chromatid segregation.

**B**

? \ ?	<b><i>A1A2</i></b>	<b><i>A3A4</i></b>	<b><i>A1A3</i></b>	<b><i>A2A4</i></b>	<b><i>A1A4</i></b>	<b><i>A2A3</i></b>
<b><i>A5A6</i></b>	A1A2A5A6	A3A4A5A6	A1A3A5A6	A2A4A5A6	A1A4A5A6	A2A3A5A6
<b><i>A7A8</i></b>	A1A2A7A8	A3A4A7A8	A1A3A7A8	A2A4A7A8	A1A4A7A8	A2A3A7A8
<b><i>A5A7</i></b>	A1A2A5A7	A3A4A5A7	A1A3A5A7	A2A4A5A7	A1A4A5A7	A2A3A5A7
<b><i>A6A8</i></b>	A1A2A6A8	A3A4A6A8	A1A3A6A8	A2A4A6A8	A1A4A6A8	A2A3A6A8
<b><i>A5A8</i></b>	A1A2A5A8	A3A4A5A8	A1A3A5A8	A2A4A5A8	A1A4A5A8	A2A3A5A8
<b><i>A6A7</i></b>	A1A2A6A7	A3A4A6A7	A1A3A6A7	A2A4A6A7	A1A4A6A7	A2A3A6A7

**Table 1-2 Phenotypic segregation ratios of a dominant allele A for the crosses of tetraploid genotypes (without considering double reduction)**

Type of crosses <sup>a</sup>	Gametes ?			Gametes ?			Zygotes				PSR <sup>b</sup>	
	AA	Aa	aa	AA	Aa	aa	A <sup>4</sup>	A <sup>3</sup> a	A <sup>2</sup> a <sup>2</sup>	Aa <sup>3</sup>	a <sup>4</sup>	a:A
NxS a <sup>4</sup> x Aa <sup>3</sup>	-	-	1	-	1	1	-	-	-	1	1	1:1
NxD a <sup>4</sup> x A <sup>2</sup> a <sup>2</sup>	-	-	1	1	4	1	-	-	1	4	1	1:5
SxS Aa <sup>3</sup> x Aa <sup>3</sup>	-	1	1	-	1	1	-	-	1	2	1	1:3
SxD Aa <sup>3</sup> x A <sup>2</sup> a <sup>2</sup>	-	1	1	1	4	1	-	1	5	5	1	1:11
DxD A <sup>2</sup> a <sup>2</sup> x A <sup>2</sup> a <sup>2</sup>	1	4	1	1	4	1	1	8	18	8	1	1:35

<sup>a</sup> N=Nulliplex, S=Simplex, D=Duplex.

Other types of crosses such as NxN, NxT (T=triplex), NxQ (Q=quadraplex), SxT, SxQ, DxT, DxQ, TxT, TxQ, QxQ, which do not segregate in the progeny, are not shown in the table.

<sup>b</sup> PSR= phenotypic segregation ratio.

of four homologous chromosomes at meiosis results in a large number of possible allelic combinations at a single locus. In the most extreme case of tetrasomic inheritance, eight different alleles descending from both heterozygous parents could segregate independently in a population, resulting in 36 possible genotypic classes in the progeny. The possible allelic configurations of both disomic and tetrasomic inheritance is shown in Table 1-1 (on page 3). The phenotypic segregation ratios are variable, ranging from 1:1 to 1:35, which depend on the dosage of parental alleles considered (Table 1-2, on page 3).

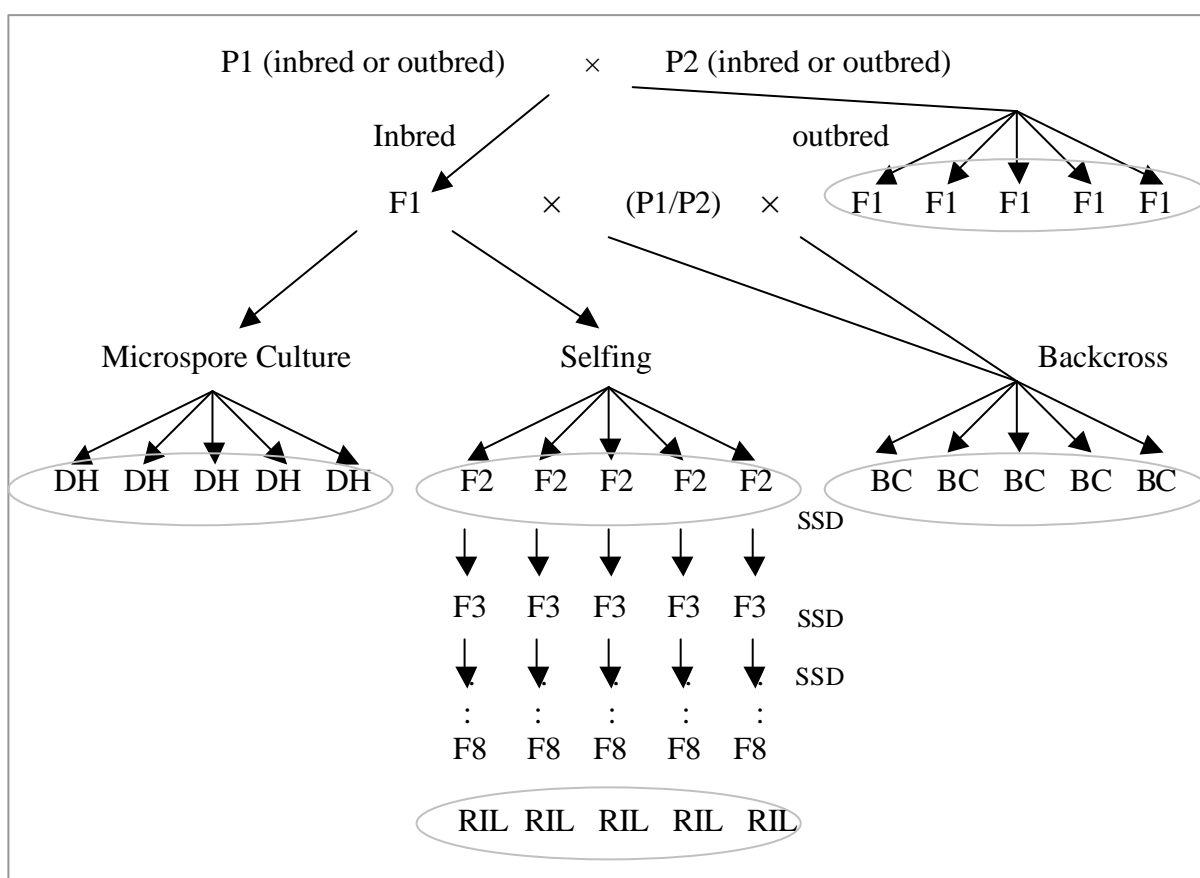
## 1.5 Molecular mapping in plants

### 1.5.1 Mapping populations

The construction of a linkage map is a process that follows the segregation of molecular markers in a segregating population and put them in linear order based on pair wise recombination frequencies . Thus, a mapping population with high number of polymorphisms over the total genome is highly desirable. Towards this end, various ways have been used to create mapping populations which are illustrated in Figure 1-2 (on page 5). Populations used for mapping are usually derived from F1 hybrids between two lines (either homozygous or heterozygous) which show allelic differences for selected probes. Specific populations can be further developed by various ways depending on the genetic features of the plant species being analysed. The commonly used populations are doubled haploid (DH) lines, recombinant inbred lines (RIL), F1, F2 and backcross (BC) populations.

Of these mapping populations, RILs offer unique advantages in QTL studies (Burr and Burr 1991, Knapp and Bridges 1990, Austin and Lee 1996). First, a well-characterized RIL population can be permanently propagated and used indefinitely without further genotyping. Second, a trait value for each genotype can be evaluated on several sister plants, which minimizes the environmental variation and improves the accuracy of the QTL analysis. Third, experiments can be replicated over years and environments using identical genotypes thus allowing the detection of QTL that cause genotype-environment interactions. And finally, maps constructed using RILs have a higher genetic resolution as compared to F2 populations (Burr and Burr 1991), as chromosome pairing and recombination during an extended period of inbreeding increases the probability of chromosome recombination. The disadvantage of using RILs is that the development of RILs is time consuming and not readily accessible to self-incompatible species like cultivated potato. Furthermore, the dominant effects at QTL cannot be estimated as done in F2. RILs have been used in maize, rice, Arabidopsis, wheat and oat for various studies such as map construction (Cho *et al.* 1998), QTL mapping (Frova

and Gorla 1993, Austin and Lee 1996, Lister *et al.* 1993, Alonso-Blanco *et al.* 1998) and candidate gene analysis (Faris *et al.* 1999, Kianian *et al.* 1999).



**Figure 12 Common strategies for the construction of mapping populations.** A DH mapping population is generated from pollen of F1 by anther or microspore culture followed by doubling of haploid. Recombinant inbred lines (RIL) are developed by successively selfing single F2 plants by single-seed descent (SSD) method until homozygosity is achieved in the F8 generation. A BC population is derived from the crossing of an F1 individual with one of the parents of F1.

Due to the self-incompatibility of potato at the diploid level and inbreeding depression, the generation of specific genetic stocks like inbred lines for molecular mapping is not applicable. Instead, segregating F1 populations from highly heterozygous parents can be employed for mapping. Such mapping populations, which make efficient use of existing genetic stocks, have been successfully used for linkage analysis in potato (Gebhardt *et al.* 1989, 1991). The estimation of recombination frequencies in such F1 populations has been investigated in detail by Ritter *et al.* (1990).

### 1.5.2 Molecular marker technologies for genetic mapping

Molecular marker technology has changed dramatically during the past two decades. The first molecular markers were isozyme markers which were based on the different mobility of differently charged protein with the same enzymatic function on the gel. Enzyme markers have limited genome coverage and numbers. Thus, they have subsequently been replaced by DNA-based restriction fragment length polymorphism (RFLP) markers, which reveal neutral sites of variation at the DNA sequence level and are independent of the environment and developmental stages and are not limited in number (Gebhardt *et al.* 1989). After that, the development of PCR-based markers such as RAPD (random amplified polymorphic DNA), AFLP (amplified fragment length polymorphism), and SSR (simple sequence repeats) has greatly accelerated and simplified the mapping process. Currently, the advent of single nucleotide polymorphism (SNP) as a new molecular marker technology has opened a new possibility for plant genome analysis (Cho *et al.* 1999).

#### **RFLP**

RFLP was first proposed by Botstein *et al.* (1980) as a new source of genetic markers in humans. This technique involves restriction of genomic DNA followed by electrophoretic size separation of the fragments in a gel matrix. The fragments are then transferred to a membrane by Southern blotting and hybridized with a radioisotope-labelled probe. Many sources of DNA can serve as probes, such as a small piece of genomic DNA, cDNA sequence, or specific PCR products. Sometimes, probes can be adopted from other species as heterologous probes for comparative mapping which allows comparison of genome organization and evolution between the related species (Tanksley *et al.* 1988, Bonierbale *et al.* 1988, Lagercrantz *et al.* 1996, 1998). The polymorphisms detected by RFLP rely on the specific and characteristic nucleotide sequence which is recognised and cut by restriction enzymes. Insertions or deletions between restriction sites or mutations occurring at restriction sites result in length polymorphisms of restriction fragments. The range of fragment length and number of fragments depends on different enzymes. Six-base cutter enzymes generate fragments with 4096 bp in length on average, while a four-base cutter enzymes generates, on average, 256 bp fragments assuming that all four bases are randomly distributed and in equal amounts in the genome. As four-base cutters score more nucleotides per length unit of DNA than six-base cutters (Kreitman and Aquade 1986), therefore, they are more efficient in detecting point mutations. RFLP markers are often co-dominant and, therefore, very informative. However, generating RFLP data is labour intensive and time consuming and requires a relatively large amount of DNA. RFLP maps have been developed for a number of species like maize (Helentjaris *et al.* 1986, Helentjaris 1987), tomato (Bernatzky and Tanksley 1986a, Helentjaris *et al.* 1986, Zamir and Tanksley 1988), lettuce (Landry *et al.*



1987b), rice (McCouch *et al.* 1988), pepper (Tanksley *et al.* 1988), Arabidopsis (Chang *et al.* 1988) and peanut (Halward *et al.* 1992).

In potato, the first RFLP map, which consisted of 135 RFLP markers, was constructed using tomato cDNA and genomic DNA as probes in an interspecific mapping population (Bonierbale *et al.* 1988). This comparative map exhibited high collinearity of marker order in both species. Based on the evaluation of genetic diversity of 38 potato lines of *Solanum tuberosum* by RFLP, Gebhardt *et al.* (1989) demonstrated the feasibility of constructing of a potato RFLP map using an intraspecific mapping population. This RFLP map with 1034 cM in length and 80 % coverage of the genome was developed using a backcross population derived from two of the 38 lines (Gebhardt *et al.* 1991). Further RFLP maps have been constructed using different mapping populations for different purposes in subsequent studies (Gebhardt *et al.* 1991, 1994, Schäfer-Pregl *et al.* 1998).

### **RAPD**

Random Amplified Polymorphic DNA (RAPD) (William *et al.* 1990, Welsh and McClelland 1990) is a PCR-based technique for DNA fingerprinting. The generation of RAPD markers is relying on the probability that 10 bp single primers of arbitrary nucleotide sequence occur on both strands of DNA in inverted orientation within a distance amplifiable by PCR. RAPD markers are technically easy to detect, only small amount of DNA and no prior sequence information is needed. RAPDs can rapidly be used to construct linkage maps. These advantages have allowed mapping of a wide variety of plant genomes (Devos and Gale 1992, Waugh and Powell 1992, Tingey and Del Tufo 1993). RAPD markers were also instrumental for map-based cloning of disease-resistance genes (Jones *et al.* 1994, Mindrinis *et al.* 1994, Whitham *et al.* 1994, Martin *et al.* 1991, Michelmores *et al.* 1991). However RAPD markers are inherited usually in a dominant manner, are not transferable from one population to another and are poorly reproducible between different laboratories (Penner *et al.* 1993, Jones 1997). Consequently, RAPD is largely being replaced by a more robust DNA fingerprinting technique termed amplified fragment length polymorphism (AFLP).

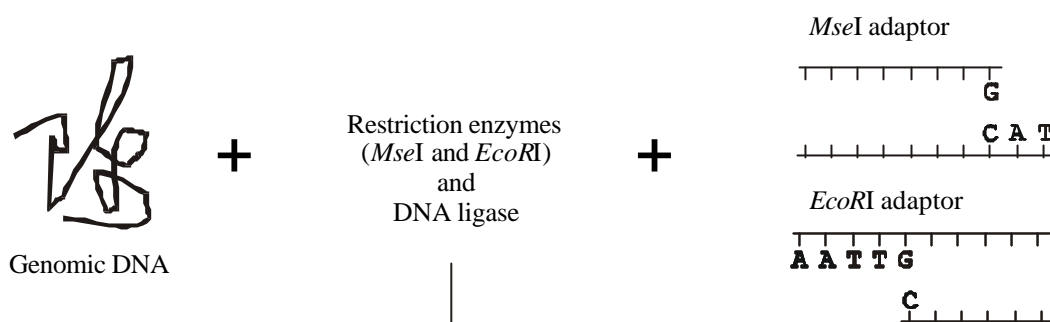
### **AFLP**

AFLP is a relatively new PCR-based DNA fingerprint technique (Vos *et al.* 1995, Zabeau and Vos 1993). It involves restriction of genomic DNA followed by ligation of adaptors to restricted fragments and preselective and selective PCR amplification of a subset of these fragments. The amplified fragments are resolved on a sequencing gel and visualised either by autoradiography or fluorescent sequencing equipment or fluorescent scanner (Meksem *et al.* 1995, Zhang *et al.* 1999, Schwarz *et al.* 1999, Huang *et al.* 2000, Huang and Sun 1999, Hartl and Seefelder 1998), depending on the method of labelling or silver

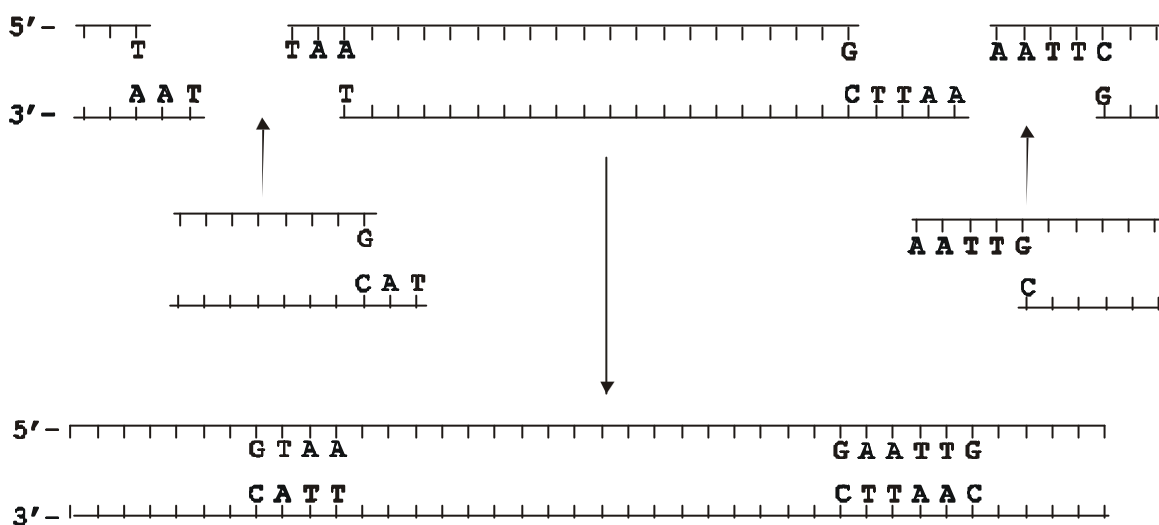
(continue to page 9)

## Introduction

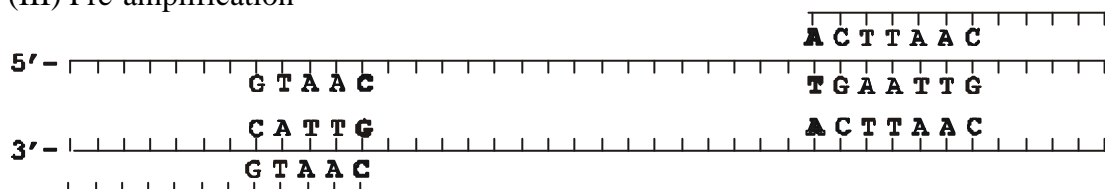
### (I) AFLP template preparation



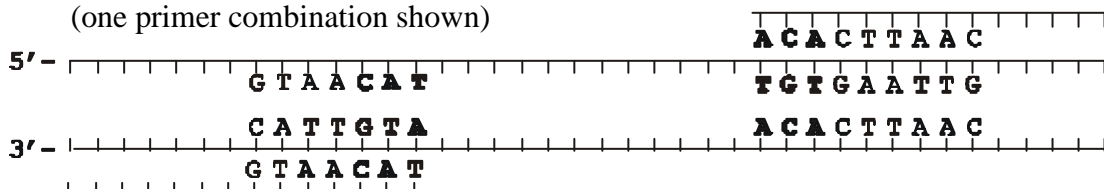
### (II) Restriction and ligation



### (III) Pre-amplification



### (IV) Selective amplification (one primer combination shown)



### (V) Electrophoresis and autoradiography or electrophoresis on 377 DNA sequencer

Figure 1-3 Schematic representation of AFLP (modified based on Mueller *et al.* 1999)

(continued from page 7)

staining (Cho *et al.* 1996). Most typically, visualisation of AFLPs is achieved by incorporating a  $^{33}\text{PO}_4^-$  group from  $[\gamma^{33}\text{-P}]\text{-ATP}$  into a dephosphorylated 5' end of one primer. This process, which involves radioactivity, was termed in the present study as radiolabelled AFLP (rAFLP). More recently, an improved AFLP technique, which was developed by Applied Biosystems (ABI), is performed by using fluorescence-labelled primers in combination with an automated sequencing system and software packages for data collection and scoring. The fluorescence-based detection of AFLPs (fAFLPs) offers a great improvement in fragment scoring and data handling. However, the few studies up to date using fAFLP as a tool are limited to genetic diversity study (Zhang *et al.* 1999), saturation mapping of a target region (Schwarz *et al.* 1999) and the chromosome assignment of fAFLP markers (Huang *et al.* 2000). The diagram summary of the AFLP technique is depicted in Figure 1-3 (on page 8).

The notable advantage of AFLP is its capacity to analyse a large number of polymorphic loci simultaneously throughout the genome with a single primer pair on a single gel without prior sequence knowledge. In contrast to RAPD, AFLP is highly reproducible and also transferable between different populations (Jones 1997, Yin *et al.* 1999, Waugh *et al.* 1997, Li *et al.* 1998, Rouppe van der Voort *et al.* 1997). AFLP has been used effectively in a variety of organisms including bacteria (Keim *et al.* 1997), fungi (Rosendahl *et al.* 1997, Van Der Lee *et al.* 1997), animals (Ajmone-Marsan *et al.* 1997, Liu *et al.* 1998) and plants. One major limitation of AFLP is the dominant nature and the difficulty in identifying allelic variants at a specific locus although codominant AFLP markers have been found, however, in frequencies of 4-15 % among all polymorphic AFLP markers (Waugh *et al.* 1997, Lu *et al.* 1998, Boivin *et al.* 1999). Furthermore, AFLP banding pattern may be affected by differential methylation of DNA in different organs from which the template DNA was extracted (Donini *et al.* 1997).

AFLP is now the first option to saturate a particular region of the genome when map-based cloning is applied to cloning target genes. AFLP has also been used in phylogenetic studies based on measures of genetic distance (Pakniyat 1997, Powell *et al.* 1996, Tohme *et al.* 1996, Russell 1997) and for identification of varieties (Schut *et al.* 1997). In addition, it has also been reported to monitor the differential expression of genes using cDNA from different plant organs (Bachem *et al.* 1996, Hartings 1999). Above all, one major application of AFLP is for molecular genetic mapping. It has been used to construct maps for barley (Becker *et al.* 1995, Waugh *et al.* 1997, Qi *et al.* 1998), sugar beet (Schondelmaier *et al.* 1996), soybean (Keim *et al.* 1997), petunia (Gerats *et al.* 1995), rice (Maheswaran *et al.* 1997) and tomato (Haanstra *et al.* 1999).

In potato, using 6 primer combinations of *EcoRI/MseI* primer set and a diploid mapping population, Van Eck *et al.* (1995) constructed an AFLP map which consisted of 264 AFLP

markers and had a map length of 1170 cM. Meyer *et al.* (1998) performed the first experiment of molecular mapping in tetraploid potato using a tetraploid breeding family in combination with AFLP marker technology. The linkage map consisted of 231 maternal and 115 paternal AFLP markers with a total map length of 991 cM and 485 cM, respectively, representing approximately 25 % coverage of the tetraploid genome.

### SSR

Simple sequence repeats (SSR), also termed microsatellite or short tandem repeats (STR), are di-, tri-, tetra- or pentanucleotide tandem repeats that are dispersed throughout the genome (Tautz 1989, Weber and May 1989). The basis of polymorphism lies in the variation of the number of tandemly repeated short nucleotide motifs which can be detected by PCR amplification (Saiki *et al.* 1988) using sequences flanking the repeats as primers. Microsatellite markers are inherited in a Mendelian fashion, they are codominant hypervariable (Saghai Maroof *et al.* 1994) and transferable within species from which primers were developed. SSRs have been used for the construction of linkage maps in *Arabidopsis* (Bell and Ecker, 1994), soybean (Akkaya *et al.* 1995), and maize (Senior *et al.* 1996). SSRs have also been used to study the allelic profiles of genotypes for purposes of genotype identification in potato (Kawchuk *et al.* 1996, Schneider and Douches 1997), soybean (Morgante and Olivieri 1993, Cregan *et al.* 1994, Rongwen *et al.* 1995, Maughan *et al.* 1995), grape (Thomas and Scott 1993) and rapeseed (Kresovich *et al.* 1995)

### SNP

SNPs, single nucleotide polymorphisms, are single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in normal individuals in natural populations (Brookes 1999). Significant efforts towards large-scale characterisation of SNPs were first initiated in human genome research. SNPs have since been shown to be the most common type of genetic variation in organisms. Of all the different types of sequence change including single nucleotide substitutions, insertions/deletions and copy number variation in nucleotide repeat motifs, SNP represents about 90 % of human DNA polymorphism (Collins *et al.* 1998). SNPs have been found to occur with a frequency as high as 1 in every 202 bp in the mouse genome and 1 in 1000 bp in the human genome (Lindblad-Toh *et al.* 2000, Wang *et al.* 1998).

There are various methods for SNP detection and scoring. More commonly used are gel-based sequencing and high-density variation-detection DNA chips (VDAs) (Wang *et al.* 1998). In the gel-based sequencing method, sequence trace data are compared using the computer program *Phred* for base calling, *Phrap* for sequence assembly and *Consed* for sequence assembly editing (Nickerson *et al.* 1997). The VDA technique involves

hybridisation to high-density DNA probe arrays or DNA chips (Wang *et al.* 1998). In addition, public EST databases can be used directly to identify SNPs by comparing sequences from different sources (Picoult-Newberg *et al.* 1999). Similarly, various techniques have been invented for genotyping SNP on large scale. The most common ones are 1) microarrays (chips) by hybridisation to allele-specific oligos (Hacia *et al.* 1999), 2) MALDI-TOF (matrix-assisted laser desorption ionisation mass spectrometry with time of flight) by detection of different molecular masses caused by the SNPs (Haff and Smirnov 1997), 3) minisequencing/single nucleotide primer extension by incorporation of specific fluorescence-labelled ddNTPs at the SNP site (Syvanen 1999). Moreover, single strand conformation polymorphism (SSCP) (Orita *et al.* 1989) and heteroduplex (HD) (White *et al.* 1992), both of which are gel-based techniques to detect conformational polymorphisms of DNA, can also be used to detect SNPs with high sensitivity (Ganguly *et al.* 1993, Sheffield *et al.* 1993) and specific mutated nucleotides showing as SNPs can be readily determined by sequencing (Schneider *et al.* 2000). In addition, the allele-specific SCAR marker is actually the gel-based detection of SNPs at sites where primers were designed for detecting specific alleles. This has been used for molecular marker conversion and mapping (Van Heusden *et al.* 1999).

SNP has many advantages and great potential for many applications. Since SNPs exist over the whole genome of organisms with a relatively high frequency, they could facilitate the development of highly dense genetic function maps that would be highly valuable for genome analysis. Moreover, as the sequence context of the SNPs is already known, it has the potential for automation and can facilitate the genetic study of associating sequence variations with heritable phenotypes on a large scale. Because of this, there has recently been considerable interest in SNP discovery and detection for genome analysis of plant (Cho *et al.* 1999)

### 1.5.3 Applications of molecular genetic map

The invention of molecular marker technology such as RFLP, RAPD, AFLP, and SSR as outlined in the previous sections has opened up a new era for genetic analysis of plant genomes. Genetic mapping using molecular marker technology is of great significance to plant breeding, plant genetic and evolutionary studies. The most common applications of genetic linkage maps are concentrated on the following areas. First, genetic linkage map can be used for marker-assisted selection (MAS) in plant breeding. Linkage maps could help identification of DNA markers linked to single genes of major agronomic importance and the tightly linked DNA markers can be used as diagnostic tools for MAS. This is particularly suitable and powerful for screening for monogenic disease resistance. One of the successful examples is MAS for soybean cyst nematode resistance (SCN) (Cregan *et al.* 1999). The SSR marker Satt309, which is located 1-2 cM away from the gene *rhg1* for resistance to SCN, has

been developed and used for tagging and tracking the gene through breeding programs, leading to the development of resistant lines. The use of SSR markers has largely decreased the time and effort involved as compared to phenotypic selection. Second, genetic linkage maps can be used for the genetic analysis of quantitative traits (see Sect. 1.7.1.1). With the construction of molecular linkage maps, characterization of quantitative traits has been greatly facilitated in identifying the genomic regions responsible for the traits, postulating the types of gene action and determining the role of epistatic effects and estimating the possible number of genetic factors controlling the traits of interest (Tanksley 1993). Third, genetic linkage mapping can be used to correlate the phenotypic traits with the genes controlling the trait, which includes map-based cloning of a gene of known heritable phenotype and postulating candidate genes for a trait with known biochemical basis. Finally, genetic linkage maps provide insights into chromosomal organization and could be useful in map-based evolutionary studies by comparative mapping.

With respect to potato, a variety of studies have been accomplished using potato molecular mapping as a tool. Many genes controlling monogenic traits have been positioned on to the potato linkage map. The majority are single genes for resistance to various pathogens. *Rx1*, *Rx2*, *Nb* and *Rx*, which condition resistance to potato virus X (PVX), were mapped on linkage groups XII and V (Ritter *et al.* 1991, De Jong *et al.* 1997, Bendahmane *et al.* 1997). *Ry* for resistance to potato virus Y (PVY) were on linkage group XI (Brigneti *et al.* 1997, Hamalainen *et al.* 1997). Five out of 11 known single dominant factors conferring race-specific resistance to *P. infestans* have been localised, *R1* on linkage group V (Leonards-Schippers *et al.* 1992, El-Kharbotly *et al.* 1994), *R3*, *R6* and *R7* on linkage group XI (El-Kharbotly *et al.* 1994, 1996a) and *R2* on linkage group IV (Li *et al.* 1998). *Gro1*, *H1* and *GroVI* conferring resistance to *Globodera rostochiensis* were mapped on to linkage group VII and V (Barone *et al.* 1990, Gebhardt *et al.* 1993, Jacobs *et al.* 1996). The *Gpa2* gene for resistance to *Globodera pallida* was mapped on to linkage group XII (Roupe van der Voort *et al.* 1997). Some tuber morphological traits were also mapped including tuber flesh colour *Y* on linkage group III (Bonierbale *et al.* 1988), flower colour loci *D*, *F*, and *P* on linkage group II, X and XI, respectively (Van Eck *et al.* 1993), tuber shape *Ro* on linkage group X (Van Eck *et al.* 1994a), tuber skin pigmentation *PSC* on linkage group X (Van Eck *et al.* 1994b, Gebhardt *et al.* 1991), tuber storage protein *Patatin* on linkage group VIII (Ganal *et al.* 1991, Gebhardt *et al.* 1991). In addition, a number of cloned genes of known function have also been mapped (Gebhardt *et al.* 1994).

## 1.6 Function mapping

Anonymous molecular markers such as RFLP, SSR, RAPD, AFLP usually have no known function. However, candidate genes analysis for QTL uses known genes as markers and

identifies their map position on chromosomes. To this end, several approaches can be employed. Southern hybridization with a cloned gene as probe is one approach to localize the gene on molecular linkage map (Gebhardt *et al.* 1994). Alternatively, linkage maps based on expressed sequence tags (ESTs) are considered as done in rice (Kurata *et al.* 1994) and maize (Chao *et al.* 1994). This depends, however, on the availability of cloned genes and a hybridization-based marker is not suitable for marker-assisted selection (MAS) screening of plants on a large-scale.

In the past decade, intensive research in plant molecular biology has led to the constant and daily increase in the number of gene sequences becoming available in databases. For example, in potato, more than 4300 sequence entries (until writing the thesis) are present in the GenBank-EMBL sequence databases, of which at least 70 genes are directly or indirectly involved in carbohydrate metabolism and transport. These metabolic pathways are physiologically and biochemically related to agronomic traits of potato such as starch and reducing sugar content of tubers. It is thus timely to explore the possibility of using sequence sources from public databases for mapping functional genes that are of agronomic interest and, in parallel, to develop STS-based and gene-specific markers.

Although many genes have been cloned and sequenced from plants and their number increases, the number of known genes available for a given species or a given pathway is still limited. This can be compensated by the use of amplified consensus genetic markers (ACGM) (Brunel *et al.* 1999). This approach takes advantage of high conservation of genes between related species and uses consensus sequence motifs as primers to amplify analogous sequences (Schneider *et al.* 1999, Van Campenhout 2000). Another approach for complementing a set of functional genes is the use of ESTs. The putative function of expressed sequence tags (ESTs) can be deduced from sequence similarity to known genes from bacteria, animals or plants by database searching. Much effort has been spent on ESTsequencing of cDNA in *Arabidopsis*, rice and tomato (Sasaki *et al.* 1994, Delseny *et al.* 1997, Ganal *et al.* 1998).

### **1.6.1 Metabolic pathways relevant to starch and sugar content of potato tubers**

The amount of starch accumulated in mature potato tubers is the net result of a chain of biochemical processes starting from photosynthetic carbon fixation followed by synthesis of transient starch and conversion into sucrose in photosynthetically active source leaves, vascular transport of sucrose from the leaves to the developing sink tuber and starch synthesis and degradation in the tuber during the growth period. These processes are biochemically related to carbohydrate metabolism and transport which include, among others, the Calvin

cycle, the oxidative pentose phosphate pathway, glycolysis, the TCA (tricarboxylic acid) cycle, sucrose metabolism and transport, and starch metabolism.

### 1.6.1.1 The Calvin cycle

The Calvin cycle also known as the C<sub>3</sub> cycle or reductive pentose phosphate pathway (Calvin and Bassham 1962), is the second phase of the dark reactions of photosynthesis. During photosynthesis, the biological energy forms (ATP and NADPH) which were converted from solar radiation in the photosynthetic light reaction, are used for fixation of carbon dioxide. The Calvin cycle has three phases including carboxylation, reduction and regeneration of sugar intermediates. The products are triose phosphates and 3-phosphoglycerate which are either transiently converted into starch by a series of reactions in chloroplast or exported to the cytosol for further metabolism such as respiration and sucrose biosynthesis. The enzymes and reactions involved are listed in Table 1-3 (on page 15) and schematically shown in Figure 1-4 (on page 16).

### 1.6.1.2 Respiratory pathways

#### *Glycolysis*

Glycolysis provides the cell with ATP and the reducing power of NADH, the cofactors to satisfy the biosynthetic needs particularly in nongreen tissues such as the fatty acid biosynthesis in leucoplasts (Dennis *et al.* 1997). Glycolysis also provides intermediates, by oxidation of glucose to pyruvate, for other biosynthetic pathways such as the synthesis of polysaccharides and lignin which are the components of cell walls (Dennis *et al.* 1997). Two glycolytic pathways exist in plants, one in the cytosol and one in plastids. These two pathways communicate through the transporters of the inner plastid envelope (Kammerer *et al.* 1998, Dennis *et al.* 1997). The enzymes catalysing the reactions within plastids are isozymes of their cytosolic counterparts (Dennis *et al.* 1997) (see Table 1-3, Figure 1-4, on page 15-16).

#### *Oxidative pentose phosphate pathway*

The oxidative pentose phosphate pathway occurs, as glycolysis, in both the cytosol and in plastids. The function of this pathway is to generate the cofactor NADPH and intermediates for other biosynthetic pathways such as the biosynthesis of nucleic acids, amino acids, the polyphenols and lignin (Dennis *et al.* 1997). The oxidative pentose phosphate pathway consists of two parts, the oxidative irreversible section, which is catalysed by G6PDH and PGDH, producing NADPH, and a reversible section that is involved in the regeneration of



## Introduction

hexose from ribulose 5-phosphate by the action of ribulose 5-phosphate epimerase, ribose 5-phosphate isomerase, transketolase, and transaldolase. The enzymes catalysing the reactions in the cytosol are isozymes of their plastidic counterparts (see Table 1-3 and Figure 1-4).

**Table 1-3 Major known enzymes involved in carbohydrate metabolism and transport**  
(see Figure 1-4 for functions)

<b><i>Calvin Cycle</i></b>		<b><i>Pentose Phosphate Pathway</i></b>	
1	ribulose biphosphate carboxylase	34	glucose-6-phosphate dehydrogenase
2	phosphoglycerate kinase	35	gluconate-6-phosphate lactonase
3	glyceraldehyde-3-phosphate dehydrogenase	36	6-phosphogluconate dehydrogenase
4	triosephosphate isomerase	37	phosphopentoisomerase
5	plastidic aldolase	38	phosphopentoepimerase
6	chloroplast fructose 1,6-bisphosphatase	39	transketolase
7	transketolase	40	transaldolase
8	transaldolase		
9	sedoheptulose-bisphosphatase	<b><i>Sucrose Metabolism</i></b>	
10	pentose-5-phosphate 3- epimerase	41	pyrophosphate-fructose-6-phosphate-1-phosphotransferase
11	ribose-phosphate isomerase	42	fructose 1,6- bisphosphatase
12	phosphoribulokinase	43	UDP-glucose pyrophosphorylase
		44	sucrose -phosphate synthase
<b><i>Glycolysis</i></b>		45	sucrose phosphatase
13	hexokinase	46	sucrose synthase
14	phosphoglucomutase	47	soluble acid invertase
15	phosphoglucose isomerase	48	fructokinase
16	phosphofructokinase	49	sucrose transporter
17	aldolase	50	H <sup>+</sup> -ATPase
18	triosephosphate isomerase	51	apoplastic invertase
19	glyceraldehyde 3-phosphate dehydrogenase	52	glucose transporter
20	phosphoglycerate kinase	53	glucose-6-P transporter
21	phosphoglycerate mutase		
22	enolase	<b><i>Starch Metabolism</i></b>	
23	pyruvate kinase	54	ADP-glucose pyrophosphorylase
		55	soluble starch synthase
<b><i>TCA Cycle</i></b>		56	granule bound starch synthase
24	pyruvate dehydrogenase	57	4- $\alpha$ -glucanotransferase (D-enzyme)
25	citrate synthase	58	branching enzyme
26	aconitase	59	L-type starch phosphorylase
27	isocitrate dehydrogenase	60	$\alpha$ -glucan phosphorylase
28	$\alpha$ -ketoglutarate dehydrogenase	61	debranching enzyme
29	succinate thiokinase	62	$\alpha$ -amylase
30	succinate dehydrogenase	63	$\beta$ -amylase
31	fumarase	64	$\alpha$ -glucosidase
32	malate dehydrogenase	65	hexokinase
33	NAD-malic enzyme	66	phosphoglucomutase



## Introduction

---

*(continued from page 14)*

The oxidative pentose phosphate pathway is intimately linked to glycolysis. These two pathways share three common intermediates: glyceraldehydes 3-phosphate, fructose 6-phosphate and glucose 6-phosphate. The metabolic flow through either of the pathways is determined by the metabolic needs of the cell. In general, carbon flow through the oxidative pentose phosphate pathway is relatively small as compared to other respiratory pathways such as glycolysis and TCA cycle (Dennis *et al.* 1997).

### ***TCA (tricarboxylic acid) cycle***

The TCA cycle, alternatively known as the citric acid cycle or Krebs cycle, occurs in mitochondria with pyruvate and malate as the main substrates. It operates as a catabolic pathway oxidizing a major portion of carbohydrates, lipids and amino acids. It has an important role in the biosynthesis of organic acids and amino acids which are the precursors for many biosynthetic pathways. In the anabolic process, the intermediates are replenished through the synthesis of oxaloacetate from phosphoenolpyruvate by phosphoenolpyruvate carboxylase (Hill 1997). Usually, important developmental stages of the plant life cycle are accompanied by increases in the rate of respiration and are often associated with increased biosynthesis of cellular constituents (Hill 1997).

### **1.6.1.3 Sucrose metabolism**

Sucrose is the principal product of photosynthesis and the major form for transport and storage in plant. Sucrose synthesis operates in the cytosol via the pathway involving the enzymes: UGPase, sucrose phosphate synthase and sucrose phosphatase (Kruger 1997). Sucrose breakdown is catalysed by sucrose synthase and invertase. The former one catalyses the interconversion between sucrose, UTP, UDP-glucose and fructose, and the latter one catalyses the irreversible hydrolysis of sucrose to glucose and fructose (Table 1-3, Figure 1-4, on page 15-16).

### **1.6.1.4 Starch metabolism**

Potato starch is a mixture of amylose and amylopectin which exist in semi-crystalline granules in plastids (amyloplasts). Amylose contains 600-3000 1,4- $\alpha$ -glucosyl residues with a 1,6- $\alpha$ -glucosyl branch every 1000 residues. Amylopectin contains 6000-60000 glucosyl residues with an average of one 1,6- $\alpha$ -glucosyl linkage every 20 to 26 units (Kruger 1997). Starch synthesis involves AGPase, starch synthases and starch branching enzymes, which are all present in multiple forms in the plant. AGPase is a tetrameric enzyme comprising of two distinct polypeptides (Muller *et al.* 1990, Nakata *et al.* 1991), a large subunit and a small subunit, both of which are encoded by several genes that are differentially expressed in

different tissues (Kruger 1997). Storage tissues contain one or more soluble starch synthases (Abel *et al.* 1996) as well as granule-bound starch synthase (Edwards *et al.* 1995). Similarly, two distinguished classes of starch branching enzyme I and II were identified (Khoshnoodi *et al.* 1996, Larsson *et al.* 1998). The activity of branching enzymes, together with that of debranching enzymes, determines the relative proportions of amylose and amylopectin (Kruger 1997).

Starch can be degraded either hydrolytically or phosphorolytically. The former pathway involves enzymes like  $\alpha$ -amylase, which catalyses the cleavage of the 1,4- $\alpha$ -glucosyl bond resulting in the production of a mixture of glucose, maltose, maltotriose and a range of branched dextrans.  $\beta$ -amylase can also degrade starch by catalysing the removal of successive maltose units from the non-reducing end of  $\alpha$ -glucan chains. The maltose and other short maltosaccharides may be further hydrolysed to glucose by  $\alpha$ -glucosidase (maltase) and subsequently phosphorylated by hexokinase. The phosphorolytic pathway involves starch phosphorylase which produces glucose 1-phosphate by successive removal of glucosyl residues at the end of an  $\alpha$ -glucan. This enzyme can only metabolise oligosaccharides larger than maltotetraose. Further phosphorolytic cleavage may attribute to the action of a glucosyltransferase such as the Denzyme which makes more of the glucan accessible to phosphorylase by increasing the degree of polymerization of short oligosaccharides (Kruger 1997). The branch points of starch are removed by a debranching enzyme which catalyses the cleavage of 1,6- $\alpha$ -glucosyl bonds hydrolytically and produces oligosaccharides.

### 1.6.1.5 Major transporters for carbon exchange between the cytosol and the plastid

Metabolite exchange between the cytosol and plastids has been identified to be achieved by three main groups of plastidic phosphate translocators including triose phosphate/phosphate translocator (TPT) (Flügge and Heldt 1991), the phosphoenolpyruvate/phosphate translocator (PPT) and the glucose 6-phosphate/phosphate antiporter (GPT) (Kammerer *et al.* 1998).

The principal route for the exchange of carbon between the cytosol and the plastid is the phosphate translocator which exchanges triose phosphate for inorganic phosphate. Other related translocators can transport other metabolites such as hexose phosphates and phosphoenolpyruvate (Borchert *et al.*, 1993, Trimming and Emes, 1993). The TPT functions predominantly in photosynthetic tissues, GPT in root and reproductive organs, whereas, PPT appears to be ubiquitously expressed (Kammerer *et al.* 1998).

Sucrose is the major form of photoassimilates for long distance transport from photosynthetic tissues to sink organs. This process involves generally three steps. First, sucrose is exported from the mesophyll cell to the vicinity of the sieve elements in the smallest veins of the leaf (short-distance transport) mainly via the symplast or through the apoplast (Frommer and

Sonneveld, 1995). This process is thought to occur by a passive mass flow. Second, sugars are transported into the sieve elements and companion cells for sieve element loading. The movement of sugars is driven by an active transport by the sucrose transport protein (SUT), from the apoplast against a concentration gradient, into the conducting complex. This is coupled with proton transport. The symport of sucrose and protons is driven by a membrane  $H^+$ -ATPase which maintains an electrochemical proton gradient across the membrane. Finally, once inside the sieve elements, sucrose and other solutes are translocated through long-distance transport from the source to sink organs by a pressure flow mechanism.

## 1.7 QTL analysis

Many traits of agricultural and economical significance exhibit quantitative inheritance, such as yield, substrate content, plant maturity, disease resistance and stress tolerance. Quantitative traits show continuous phenotypic variation in a population resulting from the combined allelic effects of many genes and environmental conditions. The genetic loci which control quantitative traits are referred to as QTL (quantitative trait loci). QTL analysis has been a major area of genetical study for many decades. The earliest documented experiments on linkage analysis between quantitative effects and marker genotypes have been reported by Sax (1923) and Thoday (1961). However, for most of the period up to 1980, the study of quantitative traits has largely involved biometrical approaches based on means, variances, covariances of relatives and, consequently, very little is known about the biological nature of quantitative or natural variation in terms of number and location of the genes that underlie them (Fisher 1918, Wright 1934, Mather 1949, Falconer 1960). It is only during the past decade when the advent of efficient molecular marker technologies as outlined in Section 1.5.2 and specific statistical methods that it was possible to follow the segregation of quantitative traits via linked markers (Tanksley 1993) and to detect effects, numbers and map positions of QTL.

### 1.7.1 Approaches to QTL analysis

The identification of QTL for economically important traits has been achieved primarily by two approaches, either through linkage mapping to anonymous markers as outlined in Sect.1.5.2. or through association studies involving candidate genes.

#### 1.7.1.1 QTL analysis through a molecular marker approach

Most QTL analyses to date adopt the linkage mapping approach. RFLPs were initially used as markers (Beckman and Soller 1983, Patterson *et al.* 1988, Lander and Bostein 1989) followed by PCR-based markers such as RAPD, SSR, and AFLP. The molecular markers,

which cover the total genome, provided the framework map based on which the QTL could be located. The principle of QTL mapping is to associate the phenotypically evaluated trait(s) with molecular markers using statistic tools. The map locations of QTL can then be estimated by the means of highly associated markers. Typically, the detection and location of the loci underlying quantitative trait variation involves three essential steps. First, a segregation population is created and characterized with molecular markers. This usually leads to the construction of a genome-wide genetic map of the population. Second, the individuals of the same population are phenotypically evaluated for the traits under investigation. Finally, genotypic molecular marker are analysed for association with the phenotypic trait data using appropriate statistical methods. This type of QTL analysis can lead to the elucidation of QTL parameters in terms of number, position, effects and interactions between them.

### 1.7.1.2 QTL analysis through a candidate gene approach

The candidate gene approach correlates a phenotype with its underlying biochemical or physiological basis by demonstrating that candidate genes are tightly linked to the genetic locus of interest. The candidate-gene approach is a powerful and robust method. Compared to the genome wide mapping strategy, the chances of finding markers linked to putative QTL are maximized, since the selection of candidate gene markers is based on known relationships between biochemistry, physiology and the agronomic character under study. This approach has been applied successfully in various QTL analyses, such as mapping QTL for defense response to diseases in wheat (Faris *et al.* 1999, Pflieger *et al.* 1999), for resistance to corn earworm in maize (Byrne *et al.* 1996, 1998) and early growth traits in maize (Causse *et al.* 1995) and bud set and bud flush in *Populus* (Frewen *et al.* 2000). The candidate gene approach is sometimes limited by the understanding of the physiology and biochemistry of the trait of interest and by the requirement of prior identification of genes that potentially involved in the trait expression. The confirmation of the causal relationship between specific QTL and candidate genes has proven to be difficult. Molecular characterization of the QTL of interest is needed to determine whether the trait variation is caused by the alleles at the candidate gene locus or at tightly linked loci and their effects. The most direct confirmation could be accomplished by creating a series of transgenic plants with alternative QTL alleles in a common genetic background and observing the phenotypic effect of each allele. Unlike the genes controlling monogenic traits, the effect of QTL alleles are likely to be more subtle and cannot be phenotypically observable on all transgenic plants. Therefore, a large number of independent transformations is needed to provide a sufficient sample size for statistical test of the difference between transformed and control plants. A more practicable alternative is the detailed analysis in marker-generated near-isogenic lines (NILs). This approach has been used for the identification of an invertase gene (*Inv*) as a QTL controlling sugar content in tomato (Bernacchi *et al.* 1998, Fridman *et al.* 2000).

### 1.7.2 QTL analysis in potato

The development of molecular linkage maps facilitates the dissection of complex traits into discrete quantitative trait loci or QTL. Thus, quantitative traits could be investigated with respect to number, effect and location of the underlying genetic factors. In potato, a number of agronomically important quantitative traits have been studied by this method such as potato tuberization, tuber shape, tuber dormancy (Van den Berg *et al.* 1996a, b, Simko *et al.* 1997) and specific gravity (Freyre and Douches 1994). One of the most important aspects in potato genetics and breeding is on quantitative disease resistance particularly to late blight. QTL for resistance to late blight (*P. infestans*) have been detected on eleven out of twelve of the potato linkage groups (Bonierbale *et al.* 1994, Leonards-Schippers *et al.* 1994, Oberhagemann *et al.* 1999). Other quantitative traits that are of much significance in potato breeding are tuber starch content, tuber yield, chip colour, and plant maturity. These traits have also been addressed in several QTL mapping experiments. Using two diploid populations, K31 and LH, and RFLP markers, Schäfer-Pregl *et al.* (1998) identified 17 QTL for tuber starch content which were distributed on all of the 12 linkage groups of potato. Also, using diploid segregating potato, Collins *et al.* (1999) and Oberhagemann *et al.* (1999) identified QTL for foliage earliness with a major QTL on linkage group V. Six QTL for chip colour, which reflected the reducing sugar content of tubers, were detected on linkage group II, IV, V, and X (Douches and Freyre 1994).

However, all these experiments outlined above were carried out at the diploid level to avoid the complexity of tetrasomic inheritance of tetraploid potato. The only mapping experiment performed at the tetraploid level was done by Meyer *et al.* (1998) and Bradshaw *et al.* (1998) in analysing QTL for resistance to late blight and *Globodera pallida*. No other published data has been reported using a tetraploid mapping population to identify QTL for agronomically important traits.

## 1.8 Agronomic traits under investigation

As outlined above, many agronomic traits of potato both qualitative and quantitative have been extensively studied regarding their nature of inheritance and map positions using diploid potato as experimental materials. In the present study particular efforts have been spent, in collaboration with potato breeding company, on the analysis of three traits: tuber starch content (TSC), reducing sugar content (RSC) and foliage earliness (FE).

### *Tuber starch content (TSC)*

Potato tubers are the agronomic product of potato cultivation while starch is the major component of the potato tubers (Sect. 1.1). High starch yield is an important target, therefore,

in potato breeding. This could be achieved by improvement of both tuber starch content and tuber yield. The present study is more focused on the former component.

### ***Reducing sugar content (RSC)***

Reducing sugar content is of critical importance to processing quality of potato. High reducing sugar levels result in dark colored chips with an undesirable taste. This is due to a non-enzymic reaction known as the Maillard reaction by which the carbonyl group of the monosaccharide reducing sugars (glucose and fructose) react with the amino groups of free amino acids during the exposure to heat in processing (Habib and Brown 1959, Shallenberger *et al.* 1959). Thus, low reducing sugar content in mature tubers is required for potato cultivars used by the processing industry.

### ***Foliage earliness (FE)***

Early maturity is a prerequisite in Middle and Northern Europe where the growing period is limited. This character is determined by different phenotypic aspects such as earliness of flowering time, tuber initiation and maturity. In potato breeding, earliness is, in general, visually evaluated as foliage earliness based on the vegetative growth of a plant.

## **1.9 Outline of the thesis**

The aim of this thesis was, first, to develop a molecular function map for potato using the diploid potato as a model system with emphasis on functional genes involved in carbohydrate metabolism and transport. The molecular function map was to be used for the identification of candidate genes controlling tuber starch content for which QTL have been mapped previously (Shäfer-Pregl *et al.* 1998). Second, to explore the possibility of discovering single nucleotide polymorphisms (SNPs) in candidate genes by sequencing directly PCR products. Third, to tag QTL for tuber starch content, reducing sugar content and foliage earliness in a tetraploid breeding family through the fAFLP approach and to clone the most significant markers via rAFLP. And fourth, to analyse the QTL through a candidate gene approach based on the function map. To achieve these goals, various experiments were performed:

Nucleotide sequence databases EMBL and GenBank were mined for sequence information of genes of potato and other plant species. Exploration was focused on genes involved in carbohydrate metabolism and transport. A sequence-tagged site (STS) approach was adopted to amplify by PCR homologous or analogous genes in potato. Polymorphisms of the amplification products were revealed either directly as SCAR (sequence characterized amplification region) or as CAPS (cleaved amplified polymorphic sequence) after cleavage with restriction enzymes in two diploid mapping populations K31 and LH. A function map containing 56 functional genes was constructed in conjunction with a previously generated



## Introduction

---

RFLP marker framework. The comparison of map positions of the functional genes with position of previously mapped QTL using the same populations lead to the identification of putative candidate genes responsible for the phenotypic variation of tuber starch content

SNPs in the candidate gene *Stp23* were exploited by sequencing PCR products from 14 different potato lines.

QTL tagging by fAFLP was performed using a fluorescent sequencing system. Cloning of significant fAFLP fragments was carried out via rAFLP.

Five genes, which are overlapped with QTL for starch content on linkage map of K31, were selected to test their effects on traits investigated in population Z3.

## 2. MATERIALS AND METHODS

---

### 2.1 Chemicals, enzymes and oligonucleotides

The restriction enzymes were obtained from Boehringer-Mannheim (10 U/μl) currently renamed as Roche (Mannheim), Biolab (England), the *Taq*-Polymerase were purchased from Life Technologies (Freiburg) or Sigma (Deisenhofen) and used with the 10 × buffer supplied, unless otherwise stated.

Laboratory reagents were obtained from following companies: Life Technologies, Pharmacia (Freiburg), Sigma (Deisenhofen), Merck (Darmstadt), Difco (USA), Roth (Karlsruhe), Serva (Heidelberg), unless otherwise stated. Filter membranes were obtained from Whatman. All types of kit were from Qiagen (Hilden) including gel-extraction kit (Qiaex II, Qiaquick). Plant genomic DNA extraction kit (DNeasy Plant Mini) and plasmid Mini and Midi Kit (QiaPrep). dNTP were from Pharmacia. Radioisotopes [ $\alpha^{32}\text{P}$ ]-dCTP (10 μCi/μl) and [ $\gamma^{33}\text{P}$ ]-ATP (10 μCi/μl) were purchased from Amersham Buchler (Braunschweig).

Oligonucleotides were custom synthesized by MWG-Biotech (Munich), Life Technologies (Freiburg) and ABI-PE (Weiterstadt).

### 2.2 Plant material

#### *Diploid Mapping populations: K31 and LH*

K31 consisting of 157 lines is an outbred diploid F1 population developed at the Max-Planck-Institute in Cologne from a cross between diploid lines: P3 (H80.557/1, *Solanum tuberosum* ssp. *tuberosum*.) as female parent and P38 (H80.576/16, *Solanum tuberosum* ssp. *tuberosum*.) as male parent (Gebhardt *et al.* 1989).

LH was derived from crossing a diploid interspecific hybrid between T710 (*S. chacoense*) as female parent and 45c3 (*S. tuberosum*) as male parent. This population consists of 49 individuals with 24 being high-starch genotypes (tuber starch content 23 % - 33 % on the basis of fresh weight) and 25 low-starch genotypes (11 % - 15 %) which have been selected from 451 F1 hybrids. For both of the populations K31 and LH, a RFLP map has been constructed and QTL analysis for tuber starch content and tuber yield has been performed (Schäfer-Pregl *et al.* 1998).

### ***Tetraploid hybrid family: Z3***

Z3 was a tetraploid F1 population consisting of 200 individuals which was derived from the cross between Elles (S3) and 90-313-1 (S1). Population Z3 was provided and developed by SaKa-Ragis Pflanzenzucht GbR (Dr. Josef Strahwald, Windeby).

### ***Potato lines for detecting single nucleotide polymorphisms (SNPs)***

Diploid lines: P3, P38, 45C3, T710.

Tetraploid lines: S1, S2, S3, Impala-L, Impala-S, Prior, Bodenkraft, Hansa, Quarta and Taiga.

## **2.3 Bacterial strain and cloning vector**

The genotype of *Escherichia coli* strain TG2: *supE hsdD5 thi D(lac-proAB) D(srl-recA)306::Tn10(tet<sup>r</sup>) F'[traD36 proAB<sup>+</sup> lac<sup>I</sup> lacZDM15]* (Life Technologies).

The *E. coli* plasmid vector pGEM<sup>®</sup>-T that contains *pMB1*-ori, *amp<sup>r</sup>* and is present in 300-400 copies per cell was used for cloning. This system provides blue and white selection for the transformant because the insertion of the cloned DNA interrupts the *lacZ* gene. The vector also contains the T7 and SP6 promoters that can serve as a priming site for insert amplification.

## **2.4 Media**

### ***Luria-Bertani (LB)***

10 g/l tryptone  
5 g/l yeast extract  
5-10 g/l NaCl  
(15 g/l agar) for LB plating

Add distilled water and adjust pH to 7.0 with NaOH, bring to 1 liter with deionised water. Autoclave at 120°C for 20 min. For plating, 15 g agar was added before autoclaving. After autoclaving, cool to 50°C, and then pour the plates either directly without adding ampicillin, or alternatively, with ampicillin added to a final concentration of 100 µg/ml.

### ***SOC medium (100 ml)***

2.0 g tryptone  
0.5 g yeast extract  
1 ml 1 M NaCl  
0.25 ml 1 M KCl  
1 ml 2 M Mg<sup>2+</sup> stock solution  
1 ml 2 M glucose

Add tryptone yeast extract, NaCl and KCl to 97 ml distilled water, stir to dissolve. Autoclave and cool to room temperature. Add filter sterilised Mg<sup>2+</sup> stock solution and filter sterilised glucose, each to a final concentration of 20 mM. Bring to 100 ml with sterile distilled water.

## Materials and methods

---

### ***2 M Mg<sup>2+</sup> stock solution (100 ml)***

20.33 g MgCl<sub>2</sub>•6H<sub>2</sub>O  
24.65 g MgSO<sub>4</sub>•7H<sub>2</sub>O

Add distilled water to 100ml, filter sterilise.

### ***White and blue selection medium (1 liter)***

1000 ml LB with 15 g agar  
2 ml 50 mg/ml Ampicillin  
2 ml 20 mg/ml X-gal  
40 µl 200 mg/ml IPTG

Make the LB plating medium as above, supplement with ampicillin, X-gal and IPTG when the medium is cooled to 50°C, and then pour the plates.

## **2.5 Buffers and solutions**

### ***50 × TAE***

2 M Tris-HCl  
50 mM EDTA  
57.1 ml glacial acetic acid

Adding distilled water to 1 liter,  
adjust pH to 8.0

### ***TE<sub>10/0.1</sub>***

10 mM Tris-HCl  
0.1 mM EDTA solution pH 8.0

### ***10 × TBE for RFLP***

890 mM Tris-HCl  
890 mM boric acid  
25 mM EDTA solution pH 8.0

### ***10 × TBE for rAFLP***

1 M Tris-HCl  
1 M boric acid  
20 mM EDTA  
solution pH 8.3 (*without adjusting*)

0.5 × TBE buffer was used  
as running buffer for rAFLP  
polyacrylamide-gel  
electrophoresis.

### ***30% Acrylamide***

29.2 % (w/v) acrylamide,  
0.8 % (w/v) N-N'-methylene bisacrylamide  
*dissolved in deionised water*

### ***rAFLP gel solution***

5 % acrylamide/bisacrylamide (20:1)  
7.5 M urea  
0.5 × TBE

***20 × TPE buffer*** (5 liters)

160 mM Tris (96.5 g)  
4 mM EDTA (7.5 g)  
pH 8.3 (*with phosphoric acid*)

0.4 × TPE was used  
for electro-blotting

***20 × SSC***

3 M NaCl  
300 mM sodium citrate

***20 × SSPE***

200 mM disodium hydrogen phosphate  
20 mM sodium dihydrogen phosphate  
3.6 M NaCl  
20 mM EDTA pH 8.0

## 2.6 DNA extraction

### 2.6.1 Plant genomic DNA extraction

#### *Plant growing and leave material preparation*

For each of the Z3 individuals, a single tuber was planted in a 2-liter pot in the greenhouse under normal daylight conditions. 10 g potato leaves are harvested, frozen in liquid nitrogen, freeze-dried and then stored at −20 °C for later use.

#### *Genomic DNA extraction*

10 mg freeze-dried leaf material was grinded under liquid nitrogen to fine powder. Total genomic DNA was extracted using DNeasy™ Plant Mini Kit (Qiagen) following the supplier's instructions. Extracted DNA was quantified as described in Sect. 2.6.5.

### 2.6.2 DNA extraction from agarose gel

PCR fragments, which were used for cloning or labelling as probes, were gel extracted using QIAquick Gel Extract Kit (Qiagen). The extraction procedures recommended by the supplier were followed. The concentration of extracted DNA was estimated as in Sect. 2.6.5.

### 2.6.3 DNA extraction from polyacrylamide gel

For cloning AFLP fragments, the QIAEXII Polyacrylamide Gel Extraction Kit (Qiagen) was used to extract DNA from polyacrylamide gel. The DNA fragment were excised from the polyacrylamide gel and crushed in 50 µl diffusion buffer followed by incubation at 50 °C for

### *Diffusion buffer*

0.5 M  $\text{NH}_4\text{Ac}$   
10 mM  $\text{MgAc}_2$   
1 mM EDTA pH 8.0  
0.1 % SDS

30 min or overnight. After centrifugation for 3 min, the supernatant was carefully transferred to a new Eppendorf tube to which 150  $\mu\text{l}$  of Buffer QX1 was added. The 10  $\mu\text{l}$  of QIAEXII were added subsequently after resuspending. DNA bound to the QIAEXII particles was eluted in 20  $\mu\text{l}$  Buffer EB (10 mM Tris-HCl, pH 8.5) after washing with PE buffer and air-drying. 2-5  $\mu\text{l}$  of the eluate was taken as template for PCR reaction in 30  $\mu\text{l}$  volume.

### 2.6.4 Plasmid DNA extraction

5 ml overnight-cultured bacteria were harvested by centrifugation and plasmid DNA was extracted using QIAprep<sup>®</sup> Spin Miniprep Kit following the supplier's protocol. The concentration of plasmid DNA was measured as in Sect. 2.6.5.

### 2.6.5 Concentration measurement of DNA

The quality and quantity of DNA were measured by comparison of band-intensity on ethidium bromide stained agarose gels with a DNA molecular weight standard, or alternatively, by the absorbance at 260 nm.

## 2.7 Cloning of PCR and AFLP fragments

### 2.7.1 Ligation of fragments in pGEM<sup>®</sup>-T

#### *A-tailing of PCR product*

The pGEM<sup>®</sup>-T needs an A-tailed PCR product for ligation. When using fresh PCR products, tailing was not necessary as *Taq* DNA polymerase has the function of adding adenosine at the end. Otherwise, an A-tailing was done before ligation.

#### *Tailing reaction*

1-2  $\mu\text{l}$  purified PCR product  
1  $\mu\text{l}$  10  $\times$  PCR buffer for *Taq* DNA polymerase  
0.5  $\mu\text{l}$  50 mM  $\text{MgCl}_2$   
1  $\mu\text{l}$  2 mM dATP  
1  $\mu\text{l}$  *Taq* DNA polymerase (5 U/ $\mu\text{l}$ )  
*add deionized water to 10  $\mu\text{l}$*   
*incubate at 70 °C for 15-30 min*

#### *Insert : vector molar ratios*

A molar ratio 3:1 of insert : vector was adopted to calculate the amount of PCR product for

## Materials and methods

---

the ligation reaction according to the following formula:

PCR product (ng) =  $3 \times (50 \text{ ng vector}) \times (\text{size of PCR product in kb}) / 3.0 \text{ kb vector}$

Ligation was done according to the supplier's instruction.

### *Ligation reaction*

5 $\mu$ l 2 $\times$ rapid ligation buffer for T4 DNA ligase
1 $\mu$ l pGEM <sup>®</sup> -T vector (50 ng/ $\mu$ l)
3 $\mu$ l PCR product (concentration optimised)
1 $\mu$ l T4 DNA ligase (3 U/ $\mu$ l)

Incubate the reaction overnight at 4 °C before transformation.

## 2.7.2 Transformation of *E. coli*

### *Preparation of competent host cells*

Glycerol stock of *E. coli* strain TG2 was streaked on LB medium and incubated at 37 °C overnight. A single colony was picked and cultured overnight in 5 ml of LB liquid medium followed by subculture of 2 ml in 500ml LB in a 2.5-liter flask. When optical density (OD) was in the range of 0.5-0.6 OD, cells were harvested by centrifugation at 5000 rpm, 4 °C for 10 min. Cells were washed 3 times by gently resuspending and centrifuging in sterile ice-cold water. Cells were then washed once using 10 ml 10 % glycerol. The cell pellet was gently resuspended in 2 ml of 10 % glycerol. The cell suspension was divided into 100  $\mu$ l -aliquot, and frozen in liquid nitrogen and stored at -70 °C.

### *Transformation*

For each transformation, 30  $\mu$ l of competent cells of *E. coli* strain TG2, thawed on ice, were mixed with 1.6  $\mu$ l of ligation mixture and transferred to the electroporation cuvette avoiding bubbles. Electroporation was done at 1.7 kV on an electroporator (BioRad). The cells were removed from the chamber in 300  $\mu$ l SOC medium, transferred to a 5 ml culture tube and immediately incubated at 37 °C for 40-90 min. 100  $\mu$ l of bacteria was plated on the white and blue selective LB media containing ampicillin (100  $\mu$ g/ml), IPTG (8  $\mu$ g/ml) and X-Gal (40  $\mu$ g/ml). Petri-dishes were incubated at 37 °C for 14-20 h.

### *Screening*

Single white colonies was picked out with a sterile tip and dipped in a PCR master mix to amplify the target insert of plasmid using appropriate primer sets. In parallel, the single colony was also streaked on an LB plate. Once the correct PCR product was obtained, the corresponding colony was propagated for plasmid DNA isolation and sequencing.

### 2.8 DNA sequencing and sequence analysis

PCR products were separated on a 1.5 % agarose gel and single bands were cut from the gel. DNA was recovered using the QIAquick Gel Extraction kit (Qiagen). Alternatively, PCR fragments were cloned into pGEM<sup>®</sup>-T vector by transforming *E. coli* strain TG2. Plasmid DNA was extracted using QIAprep Spin Mini-prep Kit (Qiagen) (see Sect. 2.6.4). The ABI PRISM Dye Terminator cycle Sequencing Ready Reaction Kit was used (Applied Biosystems) for sequencing on an Applied Biosystems 377 DNA Sequencer (Applied Biosystems).

Sequence analyses were performed using sequence analysis software package GCG (Wisconsin University, Version 9.1, UNIX, September 1997). Updated versions of the sequence databases GenBank and EMBL were searched for the sequences of functional genes available in the databases. For sequence comparison, the *Bestfit* or *Pileup* programs were used.

### 2.9 Restriction fragment length polymorphism (RFLP) analysis

#### 2.9.1 Genomic DNA restriction and purification

5 µg genomic DNA were digested overnight by 20 U *TaqI* or *AluI* in buffer B at 65°C or buffer A at 37°C. Digested DNA was mixed with 2 volumes of ice-cold 100 % ethanol and 1/20 volume of 4 M NH<sub>4</sub>Ac, precipitated and collected by centrifugation at 14000 rpm for 30 min. The pellet was washed with 70 % ethanol and air-dried, followed by adding 5 µl formamide loading buffer. The pellet was resuspended by vortexing, denatured at 95 °C for 5 min and chilled on ice before loading.

#### 2.9.2 Fragment separation by electrophoresis

DNA fragments were separated on a 4 % denaturing polyacrylamide gel as described by Kreitman and Aquade (1986) and modified by Gebhardt *et al.* (1989).

##### *Gel mix*

62.4 g urea 13 ml 40 % polyacrylamide/ bisacrylamide (20:1) 13 ml 10 × TBE <i>add deionized water to 130 ml</i>
--

Gel mix was filtered through 0.2 µm filter and degassed under vacuum for 5 min. After adding 200 µl 25 % APS (ammonium

persulfate) and 150 µl TEMED and gentle mixing, the gel was poured. The gel was 30 cm wide, 40 cm high and 0.1 cm thick with loading slots of 0.4 cm in width. Gel was polymerized for at least two hours at room temperature before loading the samples. The



## Materials and methods

---

electrophoresis was run in  $1 \times$  TBE (1.8 liter) (see Sect. 2.5) as running buffer. Power was set first to 15 W for 15 min followed by a 5-6 h run at a constant power of 50 W for until the xylene cyanol marker reached the bottom of the gel.

### ***Formamide loading buffer***

94 % formamide  
0.05 % xylene cyanol  
0.05 % bromo phenol blue  
10 mM EDTA pH 7.2

### **2.9.3 Electro-blotting of DNA fragments**

Electro-transfer of DNA fragments onto Nylon-membrane (Pall, Biodyne) was performed in 23 liters of  $0.4 \times$  TPE buffer in a home-made blotting chamber (45 cm long and 35 cm wide) for 1 hour at 35 V (5 V/cm). DNA was covalently cross-linked to the membrane by drying at 80°C for 30 min. The membrane was then used for hybridisation or stored in  $2 \times$  SSC solution at 4 °C.

### **2.9.4 Radioactive labelling of DNA probes**

#### ***5 × OLB buffer***

0.1 mM dATP, dCTP, dGTP, dTTP  
1.0 M HEPES, pH 6.6  
2.0 mg/ml Hexadesoxyribonucleotide pdN6 (Pharmacia)  
0.2 M Tris-HCl, pH 8.0  
 $3.7 \times 10^{-5}$  M  $MgCl_2$   
0.35 % (v/v)  $\beta$ -Mercaptoethanol

#### ***Sephadex buffer***

10 mM Tris-HCl, pH 7.5  
1 mM EDTA  
0.3 mM NaCl  
0.1 % SDS

*filter sterilise or autoclave*

#### ***Sephadex suspension***

100 ml Sephadex buffer  
2 g Sephadex-G50

*autoclave*

#### ***Random primer labelling reaction (20 $\mu$ l)***

4  $\mu$ l 5 × OLB  
1  $\mu$ l Klenow-Polymerase (2 U/ $\mu$ l)  
3  $\mu$ l of [ $\alpha^{32}$ P]-dCTP (10  $\mu$ Ci/ $\mu$ l)  
12  $\mu$ l (50-100 ng) denatured DNA

50-100 ng of gel-purified DNA was denatured in 12  $\mu$ l volume at 95°C for 5 min and

chilled immediately on ice. The random primer labelling method (Feinberg and Vogelstein 1983, 1984) was used to label the probe DNA with [ $\alpha^{32}$ P]-dCTP (10  $\mu$ Ci/ $\mu$ l) in a total reaction volume of 20  $\mu$ l.

Leave the labelling reaction at room temperature overnight or at 37°C for 1h. The separation of labelled DNA-fragments from unincorporated nucleotides was done using a 1 ml-Sephadex G-50 column. The 1 ml syringe was filled with sterile glass wool at the bottom and G-50 Sephadex equilibrated in buffer (G50-Sephadex suspension). The labelling reaction was mixed with 300  $\mu$ l Sephadex suspension and eluted through the column. A second elution

with the same amount of Sephadex suspension was performed and the second elution was collected. Labelling efficiency was checked using the radioactive counter. The DNA probe was mixed with 80 µl hering sperm DNA (5 mg/ml) and denatured at 95°C for 5 min.

### 2.9.5 Hybridisation

Southern blot hybridization was performed in glass tubes (30 cm x 3.5 cm in diameter). Up to 3 membranes were hybridised in one tube. The membranes were stacked on top of each other, then rolled up with a plastic stick and unrolled around the inner wall of the tube. Tubes were tightly closed with plastic screw caps with sealing rubber inside. 10 ml 6 × SSC per membrane and 40 µg/ml denatured hering sperm DNA were added to the tube for pre-hybridisation and hybridisation. For pre-hybridization the tubes were incubated at 65 °C, rotating around the long axis in a hybridisation oven (Bachofer, Reutlingen, Germany) for at least 4 hours. After adding the denatured probe, the hybridisation was performed overnight under the same conditions as pre-hybridisation.

After hybridisation the filters were washed at 65 °C in the hybridisation oven in the same tube three times 15 min each, using 50 ml of pre-heated washing solution (1 × SSC, 0.1 % SDS). Then the filters were taken out from the tube and washed two times in a tank containing sodium phosphate buffer at room temperature for at least 30 min. The filters were wrapped in thin plastic foil (Saran film) and placed in cassettes with intensifying Trimax-screen. Kodak X-OMAT AR5 film was placed on the filter and exposed at –70°C for about 10 days before developing. Alternatively, filters were exposed on Fuji phosphoimage screens for 5 days. Fragment patterns were visualised using a Fuji BAS-2000 Phosphoimager (Fuji Photo Film Company Ltd., Japan).

#### *6 × SSC*

55.2g/l NaCl  
26.4g/l Sodium citrate  
(pH7.0)

#### *Hybridising solution*

7 % SDS  
0.12 M Na<sub>2</sub>HPO<sub>4</sub>  
1 mM EDTA  
40 µg/ml denatured  
hering sperm DNA,  
(pH 7.0)

#### *Washing buffer*

1 × SSC  
0.1 % SDS

#### *Sodium phosphate buffer*

5 mM Na<sub>2</sub>HPO<sub>4</sub>  
1 mM EDTA  
0.2 % SDS, pH7.0

## 2.10 Sequence-tagged sites (STS) analysis

### 2.10.1 Exploration of functional gene sequences in databases

GenBank and EMBL databases were explored for sequence information of genes involved in carbohydrate metabolism including the Calvin cycle, the oxidative pentose phosphate

pathway, the citric acid cycle, glycolysis, sucrose and starch metabolism and intermediate transport. The exploration of functional genes was also expanded to other plant species such as tomato, tobacco and *Arabidopsis* in case when gene sequences are not available in potato.

### 2.10.2 Nomenclature of the functional genes

The rules of CPGN (the Commission on Plant Gene Nomenclature) were followed for gene nomenclature in potato. Genetic loci and corresponding genes were given italicised three-letter designations, starting with one upper case letter. The three letters reflect a particular property of gene function. Different loci detected by the same gene were distinguished by the assignment of an additional letter a, b or c. Gene products were designated in non-italicised capital letters. When the same three-letter code has been assigned to genes in different species, the species two-letter code '*St*' for potato, '*Le*' for tomato and '*At*' for *Arabidopsis* are added to gene names.

### 2.10.3 Polymerase chain reaction (PCR)

#### Primer design

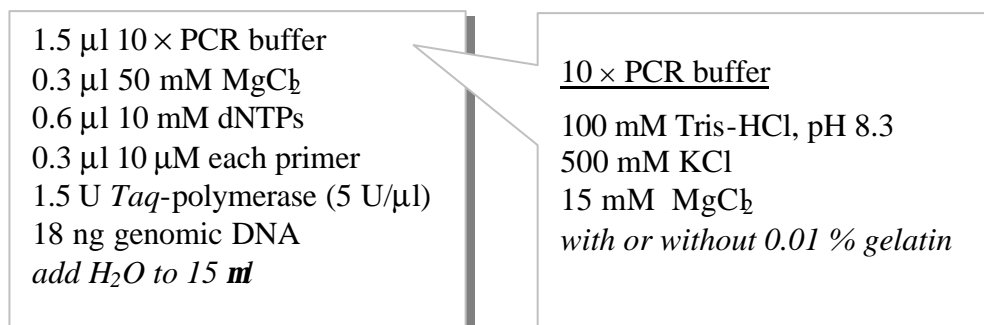
The following criteria were followed to design STS primers. First, primer design should avoid the intron-exon junctions since most of the available sequences in the public databases are cDNAs. Second, for detection of polymorphism, primer positions are ideally chosen to frame one or two introns since these regions are potentially variable at the intraspecific level (Weining and Langridge 1991; Tragoorung *et al.* 1992; Konieczny and Ausubel 1993; Cote-Real *et al.* 1994). Finally, primers should have good characteristics for amplification with optimal annealing temperature of 55-60 °C and primer length between 22 and 24 bases.

In cases when gene sequences were not available in potato, tomato sequences were employed instead. When neither a potato nor a tomato sequence was available, degenerate primers based on sequence conserved between genes from other plants were designed. Consequently, these primers were composed of a mixture of several different oligonucleotides supposedly in equivalent proportion.

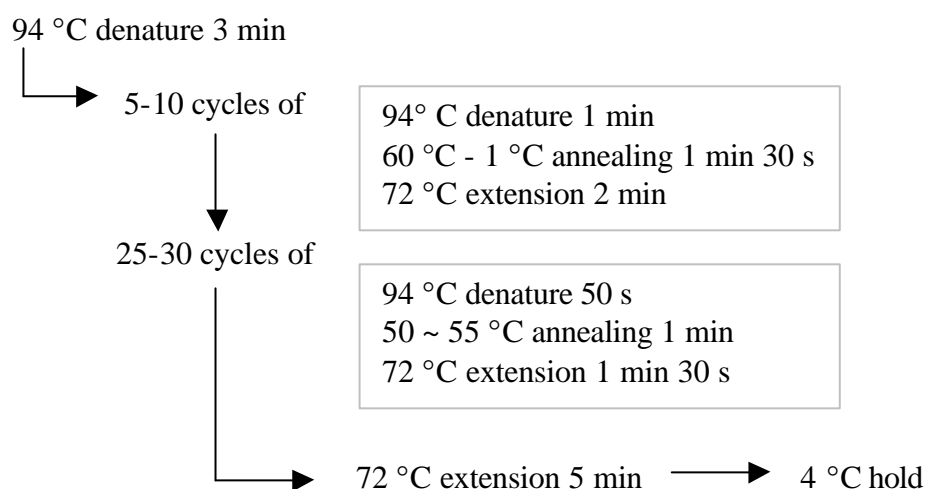
#### PCR reaction conditions

PCR reaction was performed in 15-100 µl reaction volume depending on the amount of PCR product needed. Mostly, PCR reactions were set up in a final volume of 15 µl. The reactions mix was capped with a drop of mineral oil that was omitted when reactions were performed in a heating-lid thermo-cycler (Biometra, Göttingen). A 'touchdown' amplification protocol was employed (Kresovich *et al.* 1995) with the following conditions:

### PCR reaction (15 ml)



### PCR conditions



#### 2.10.4 Identification of sequence characterized amplified regions (SCAR)

To discover SCAR polymorphisms, 5-10 individuals of populations K31 and LH were randomly selected for PCR amplification. PCR products were analysed by electrophoresis on 1.5 % - 2 % agarose gel in 1 $\times$  TAE. 1-kb ladder was used as size marker. DNA fragments were visualised by ethidium bromide staining.

#### 2.10.5 Identification of cleaved amplified polymorphic sequences (CAPS)

Monomorphic PCR products were digested for CAPS with a range of restriction endonucleases. These restriction enzymes were 4-bp cutter including *AluI*, *RsaI*, *TaqI*, *HpaII*, *DdeI*, *HaeIII*, *MseI*, and *MboI*. Two units of restriction enzyme, mixed with the reaction buffer, were added to 15  $\mu$ l of the PCR reaction to final volume of 20  $\mu$ l. The digests were incubated at least 3 h or overnight at the recommended temperature. CAPS polymorphisms were analysed by agarose gel electrophoresis as described above.

### 2.11 Generation of segregation data and linkage analysis

When the approaches as outlined in Sect. 2.10.4 and 2.10.5 detected a polymorphism, parents P3, P38 and 90-150 progeny of populations K31 and 49 progeny and parental lines T710 and 45C3 of population LH were analysed by PCR. Fragments were scored as 0, 1 and 2 for absence, presence and missing value, respectively. The segregation data was then added to the existing RFLP framework data sets (Schäfer-Pregl *et al.* 1998). Recombination frequencies were computed using the software package MAPRF (version 2.1, E. Ritter). The recombination frequency between any two markers was displayed in the MAPDIC sub-program. Linkage of a new marker to an existing linkage group was identified using the FIND command. The PATTERN command was used to check the recombinant genotypes for scoring errors (singletons). The genetic distance between marker pairs is measured by the recombination frequency. The linear order of the markers in a linkage group was determined based on pairwise recombination frequencies in two-, three-, or multiple-point estimates (Ritter *et al.* 1990). The recombination values were converted into map distances (cM) by Kosambi's mapping function (Kosambi 1944).

### 2.12 Detection of single nucleotide polymorphisms (SNPs)

Gene *Stp23* was selected as a case study for detecting SNPs in functional genes. Genomic DNA from 14 potato lines (see Sect. 2.2) was used to amplify the homologous segment of gene *Stp23* by PCR. PCR products were bi-directionally sequenced. The PCR products of three lines P3, P38 and T710 were additionally cloned into pGEM<sup>®</sup>-T vector prior to sequencing. Electropherograms of nucleotide sequence data were visually examined for alternative nucleotides (SNPs) at specific positions in individual genotypes. Sequence data of different lines were assembled by *Pileup* program of GCG software package (see Sect. 2.3). Each vertical slice of the resulting assembly was examined for nucleotide variation. Any such slice was considered a candidate SNP.

### 2.13 Amplified fragment length polymorphisms (AFLP)

#### 2.13.1 Radioisotope-labelled amplified fragment length polymorphism (rAFLP) analysis

##### DNA digestion and ligation of adaptors

35-250 ng of genomic DNA was incubated for 1 h at 37 °C with 5 units of *Eco*RI and 5 units of *Mse*I, 4 µl of 10 × RL buffer in a 40 µl reaction volume.

##### *Restriction reaction* (40 µl)

## Materials and methods

---

4  $\mu$ l 10  $\times$  RL buffer  
2  $\mu$ l *Eco*RI (10 U/ $\mu$ l)  
2  $\mu$ l *Mse*I (10 U/ $\mu$ l)  
35-250 ng genomic DNA  
*add deionised water to 40  $\mu$ l*

10  $\times$  RL buffer  
100 mM Tris-HCl  
100 mM MgAc  
500mM KAc

Adaptors were prepared by mixing equal-molar amounts of both stands, denatured at 95 °C for 5 min and renatured at room temperature over 10 min.

### **Ligation reaction** (50 $\mu$ l)

40  $\mu$ l restriction mixture  
1  $\mu$ l *Eco*RI-adaptors (5 pM/ $\mu$ l)  
1  $\mu$ l *Mse*I-adaptors (50 pM/ $\mu$ l)  
1  $\mu$ l 10 mM ATP  
1  $\mu$ l 10  $\times$  RL buffer  
1 unit of T4 DNA ligase

10  $\mu$ l of a solution containing 5 pMol of *Eco*RI-adaptors, 50 pMol of *Mse*I-adaptors, 1  $\mu$ l of 10 mM ATP, 1  $\mu$ l of 10  $\times$  RL buffer, and 1 unit of T4 DNA ligase was added to the restriction solution, and the incubation was continued for 3 h at 37 °C.

After ligation, the reaction mixture was diluted 1-10 fold with TE10/0.1 (see Sect. 2.5) buffer depending upon the amount of DNA used for restriction. The reference dilution was 10-fold when 250 ng genomic DNA were used for the restriction and ligation. Reactions with different amounts of DNA were diluted based on the their proportion of DNA to that of the reference dilution. The calibrated restriction and ligation mixtures were stored at -20 °C.

## **Preselective amplification**

### **Preselective amplification reaction** (20 $\mu$ l)

2  $\mu$ l 10  $\times$  PCR buffer  
0.4  $\mu$ l 50 mM MgCl<sub>2</sub>  
0.8  $\mu$ l 10 mM dNTPs  
0.4  $\mu$ l 10  $\mu$ M each primer  
0.4  $\mu$ l *Taq*-polymerase (5 U/ $\mu$ l)  
4  $\mu$ l template DNA  
*add H<sub>2</sub>O to 20  $\mu$ l*

#### *Reaction conditions*

The pre-amplification reactions were performed for 20 cycles with the following cycle profile: a 30 s DNA denaturation step at 94°C, a 30 s annealing step at 56 °C, and a 1 min extension step at 72°C. The amplification products were diluted 20-fold with TE10/0.1 buffer before proceeding to selective

amplification.

## **Labelling of primers and size marker with [<sup>33</sup>P]-ATP**

### **Primer labelling reaction** (12.5 $\mu$ l reaction volume for 25 selective amplification reactions)

The labelling reaction mix was incubated at 37 °C for 45 min, and then stopped by incubating at 75 °C for 10 min.

## Materials and methods

---

1.3  $\mu$ l 10  $\times$  T4 PNK buffer  
0.4  $\mu$ l T4 polynucleotide kinase (PNK) (6 U/ $\mu$ l)  
2.5  $\mu$ l *Eco*RI-primer (10 pM/ $\mu$ l)  
2  $\mu$ l [ $\gamma^{33}$  P]-ATP (10  $\mu$ Ci/ $\mu$ l)  
*add H<sub>2</sub>O to 13  $\mu$ l*

10  $\times$  T4 PNK buffer  
250 mM Tris-HCl pH 7.5  
100 mM MgCl<sub>2</sub>  
50 mM DTT  
5 mM Spermidine-HCl

### *Size marker labelling*

30-330 bp AFLP DNA ladder, not dephosphorylated, was radioactively labelled either by the exchange reaction or as in previous part after dephosphorylation.

### *Labelling by exchange reaction*

Set up reaction mix in 0.2 ml PCR tube:

1  $\mu$ l 5  $\times$  Exchange reaction buffer  
2  $\mu$ l 30-330bp AFLP DNA ladder  
1  $\mu$ l [ $\gamma^{33}$  P]-ATP (10  $\mu$ Ci/ $\mu$ l)  
1  $\mu$ l T4 polynucleotide kinase (10 U/ $\mu$ l).

5  $\times$  Exchange reaction buffer  
250 mM imidazole (pH 6.4)  
60 mM MgCl<sub>2</sub>  
5 mM 2-mercaptoethanol  
350  $\mu$ M ADP

Collect solution by brief centrifugation, incubation for 10 min at 37 °C and stop the reaction by heating the tube 15 min at 65 °C.

Before loading on the gel, add equal volume of loading buffer to AFLP samples.

### *Dephosphorylation of marker DNA (40 $\mu$ l)*

5  $\mu$ l 30-330 bp AFLP DNA ladder (Life Tech.)  
4  $\mu$ l 10  $\times$  SAP buffer  
2  $\mu$ l Shrimp alkaline phosphatase (1 U/ $\mu$ l)  
29  $\mu$ l H<sub>2</sub>O

Incubate at 37 °C for 45 min, stop the reaction by addition of 0.5  $\mu$ l 0.5 M EDTA (pH 8.0) and incubate at 75 °C for 10 min. The dephosphorylated DNA marker can be use directly for labelling without precipitation and purification.

### *Labelling the dephosphorylated AFLP DNA ladder*

Set up the labelling reaction in a 0.2 ml PCR tube (10  $\mu$ l). Incubate for 45 min at 37 °C,

7.5  $\mu$ l dephosphorylated AFLP DNA ladder  
1  $\mu$ l 10  $\times$  T4 PNK buffer  
1  $\mu$ l [ $\gamma^{33}$  P]-ATP (10  $\mu$ Ci/ $\mu$ l)  
0.5  $\mu$ l T4 PNK (6 U/ $\mu$ l)

followed by 10 min at 70 °C to stop the reaction. 1-2  $\mu$ l was taken and added equal volume of AFLP loading buffer before gel loading.

## Materials and methods

---

### Selective amplification

Selective amplification was carried out with two oligonucleotide primers to which two more selective nucleotides had been added at the 3' end. One of the two primers, *EcoRI*-primer, was end-labelled. PCR was performed in 20 µl reaction volume.

#### *Selective amplification reaction* (20 µl)

2 µl 10 × PCR buffer  
0.4 µl 50 mM MgCl<sub>2</sub>  
0.4 µl 10 mM dNTPs  
0.6 µl 50 ng/µl non-labelled primer  
0.5 µl labelled primer  
0.2 µl *Taq*-polymerase (5 U/µl)  
5 µl diluted preamplification products  
*add H<sub>2</sub>O to 20 µl*

#### Selective amplification reaction conditions

The AFLP reactions with primers having three selective nucleotides were performed for 36 cycles with the following cycle profile: a 30 s DNA denaturing step at 94 °C, a 30 s annealing step, and a 1 min extension step at 72 °C. The annealing temperature was 65 °C in the first cycle,

subsequently reduced by 0.7 °C per cycle for the next 12 cycles, finally stabilising at 56 °C for the remaining 23 cycles. All amplification reactions were carried out in a same thermocycler.

### Denaturing polyacrylamide gel analysis of AFLP

The AFLP selective amplification products were mixed with an equal volume (20 µl) of formamide loading buffer.

#### *Formamide loading buffer for rAFLP*

98 % formamide  
10 mM EDTA pH 8.0  
0.025 % bromophenol blue  
0.025 % xylene cyanol

The samples were denatured for 3 min at 95 °C, and then quickly chilled on ice. 4 µl of each sample were loaded on a 5 % denaturing polyacrylamide gel.

#### *The rAFLP gel mix*

75 ml 5 % acrylamide/ bisacrylamide  
(20:1) solution  
50 µl TEMED  
750 µl 10 % APS

5 % acrylamide/ bisacrylamide  
(20:1) solution (1 liter)

250 ml acrylamide-bisacrylamide (20:1)  
450 g urea  
50 ml 10 × TBE for AFLP  
*add distilled water to 1 litre*  
*filter sterilised and stored at 4 °C*

The 5 % acrylamide / bisacrylamide (20:1) solution was degassed 5 min. Then 750 µl



## Materials and methods

10 % APS (ammonium persulfate) and 50 µl TEMED were added and gently mixed. Gels were cast using a 38 × 50 cm gel glass plate apparatus (BioRad). The gels were fixed onto one of two glass plates, one had been treated with sticking solution (0.3 % acetic acid and 0.015 %  $\gamma$ -methacryloxypropyltrimethoxysilane in ethanol) whereas the other one with repellent solution (Acrylease™, Stratagene, Heidelberg, Germany). The gel was polymerized for at least two hours or overnight at room temperature. A pre-run was performed for 30 min in pre-heated 0.5 × TBE buffer before loading the gel. The gel was run at a constant power of 58 W for 2-3 h until the xylene cyanol marker reached the bottom of the gel.

After electrophoresis, the gel on the glass plate was fixed in 10 % acetic acid for 30 min, followed by rinsing with distilled water. After drying in an oven at 80 °C for 1.5 h, the glass plate was wrapped with Saran film and exposed to a Kodak XOMAT AR5 film at room temperature for 1-2 days before developing.

### 2.13.2 Fluorescent dye-labelled amplified fragment length polymorphism (fAFLP) analysis

Table 2-1 Adaptor and primer sequences and fluorescent labelling dye used for fAFLP analysis

Primer/adaptor	Labelling dye	Sequence (5'-3')
<i>EcoRI</i> -adaptor		CTCGTAGACTGCGTACC CATCTGACGCATGGTTAA
<i>EcoRI</i> +1 primer <i>EcoRI</i> +3 primers	FAM (blue) JOE (green) TAMRA (yellow)	GACTGCGTACCAATTC-a (=E) E-aca, E-act E-aag, E-acg, E-agg E-agc, E-acc
<i>MseI</i> -adaptor		GACGATGAGTCCTGAG TACTCAGGACTCAT
<i>MseI</i> +1 primer <i>MseI</i> +3 primers		GATGAGTCCTGAGTAA-c (=M) M-caa, M-cac, M-cag, M-cat, M-cta, M-ctc, M-ctg, M-ctt
<i>HindIII</i> -adaptor		CTCGTAGACTGCGTACC CATCTGACGCATGGTCGA
<i>HindIII</i> +1 primer <i>HindIII</i> +3 primer	FAM (blue) JOE (green) TAMRA (yellow)	AGACTGCGTACCAGCTT-a (=H) H-act H-agt, H-aga H-ata
<i>TaqI</i> -adaptor		GACGATTGAGTCCTGAC TAACTCAGGACTGGC
<i>TaqI</i> +1 primer		ATTGAGTCCTGACCGA-a (=T) ATTGAGTCCTGACCGA-g
<i>TaqI</i> +3 primer		T-agt, T-atg, T-gta, T-gat

## Materials and methods

---

For fAFLP genotyping with *EcoRI/MseI* primer sets, reagents including *EcoRI* adaptor pairs, *MseI* adaptors, preselective and selective amplification primers and AFLP Core Mix were obtained from Applied Biosystems (ABI). For screening with other primer sets such as *EcoRI/HindIII*, *EcoRI/TaqI* and *HindIII/TaqI*, primers and corresponding adaptors were designed and custom synthesised by MWG-Biotech (Munich). The sequences of adaptors and primers are shown in Table 2-1 (on page 39). The fAFLP procedures followed the supplier's instructions (PE Applied Biosystems).

### Restriction and ligation reaction

*Annealing of the adaptors* (see Sect. 2.13.1)

*Restriction enzyme master mix* (for 50 DNA samples in 50  $\mu$ l)

5  $\mu$ l 10  $\times$  T4 DNA ligase buffer with ATP  
5  $\mu$ l 0.5 M NaCl  
2.5  $\mu$ l 1 mg/ml BSA  
50 U *MseI*  
250 U *EcoRI*  
50 U T4 DNA ligase (400 U/ $\mu$ l)  
*add sterile deionised water to 50  $\mu$ l*

1  $\times$  T4 DNA ligase buffer with ATP  
50 mM Tris-HCl (pH 7.8)  
10 mM MgCl<sub>2</sub>  
10 mM dithiothreitol  
1 mM ATP  
25  $\mu$ g/ml bovine serum albumin

*Reaction mix* (11  $\mu$ l)

5.5  $\mu$ l (35 ng-500 ng) genomic DNA  
1  $\mu$ l 10  $\times$  T4 DNA ligase buffer with ATP  
1  $\mu$ l 0.5 M NaCl  
0.5  $\mu$ l 1 mg/ml BSA  
1  $\mu$ l *EcoRI* adaptor pair  
1  $\mu$ l *MseI* adaptor pair  
1  $\mu$ l restriction enzyme master mix

Mix thoroughly and spin down to the bottom of the tube. Incubate at 37 °C for 3 h using a thermal cycler with a heated lid to prevent *EcoRI*\* (star) activity.

### *Dilution of the restriction-ligation samples*

Dilution of restriction-ligation samples depended on the amount of DNA used. When 500 ng genomic DNA were used, the sample was diluted 20-fold by adding 189  $\mu$ l of TE<sub>10/0.1</sub>. Otherwise, dilution was adjusted described in Sect. 2.13.1.

## Materials and methods

---

### Amplification of target sequences

#### *Preselective amplification reaction*

4 µl diluted restriction-ligation DNA  
1 µl AFLP preselective primer pairs  
15 µl AFLP Core Mix\*

\* AFLP Core Mix was a reagent mixture including AmpliTaq enzyme, buffer and dNTPs that was provided by ABI.

#### *Preselective amplification reaction conditions*

The preselective amplification reactions were performed with the following PCR profile: first incubate at 72 °C for 2 min, then 20 cycles of a 1 s at 94 °C, a 30 s at 56 °C, 2 min at 72 °C, and finally, incubating at 60 °C for 30 min.

#### *Dilution of preselective amplification products*

The preselective amplification reaction products were diluted 20-fold with TE<sub>10/0.1</sub> before proceeding to the selective amplification.

#### *Selective amplification reaction*

1.5 µl diluted preselective amplification product  
0.5 µl 5 µM *Mse*I Primer  
0.5 µl 1 µM dye-labelled *Eco*RI primer\*  
7.5 µl AFLP Core Mix

\* the *Eco*RI primers were either labelled with 5-carboxy-fluorescein (5-FAM, blue), 2',7'-dimethoxy-4',5'-dichloro-6-carboxy-fluorescein (JOE, green) or N,N,N',N'-tetramethyl-6-

carboxy-rhodamin (TAMRA, yellow). The internal size standard was Genescan-500 ROX length standard (PE Applied Biosystems) labelled with 6-carboxy-X-rhodamin (ROX, red).

#### *Selective amplification reaction conditions*

First cycle: 94 °C for 2 min, 65 °C for 30 s, 72 °C for 2 min, next 8 cycles: 94 °C for 1 s, 64 °C for 30 s (with 1 °C decreased for each cycle) and 72 °C for 2 min, followed by 23 cycles of 94 °C for 1 s, 56 °C for 30 s and 72 °C for 2 min with a final step of incubation at 60 °C for 30 min.

#### *Samples preparation*

##### Loading dye mix (for 100 samples)

170 µl formamide  
40 µl blue dye (loading buffer, ABI)  
30 µl size standard of GS-500 ROX

2 µl selective amplification product of each sample was mixed with 2.4 µl aliquot of loading dye denatured at 95 °C for 2 min, then immediately chilled on ice until ready for loading.

### Gel electrophoresis on 377 DNA Sequencer

#### *Plate preparation*

Wash the plates two times with 2N NaOH and Alconox™ alternately and rinse with distilled water, stand the plates up right and allow to air dry. Similarly, wash, rinse and dry the comb and the spacers.

#### *Gel preparation*

##### Gel mix for fAFLP

21 g urea (Merck)  
8.4 ml 30 % acrylamide/bisacrylamide (29:1) solution  
6 ml 10 × TBE (Roth)  
20 ml H<sub>2</sub>O (Merck)

Stir the gel mix solution until urea is completely dissolved. Filter the solution through a 0.2 µm cellulose nitrate filter and degas for 5 min. Add 350 µl 10

% APS and 15 µl TEMED, swirl gently to mix and cast the gel by injecting the solution with a syringe in the space between the plates. Insert the comb avoiding to trap bubbles at the interface. Allow gel to polymerise for 2 h before use.

#### *Instrument set up*

Clean read region and comb area of the plates with Kimwipes™ and distilled water, then place the gel cassette against the heat transfer plate of the instrument. Fill the upper buffer chamber with 580-600 ml 1 × TBE (Roth) and check for leaks. Fill the lower buffer chamber with 1 × TBE until the bottom of the glass plates is covered. Place the front heat-transfer plate on the front of the cassette. Secure it with and attach the quick-connect water lines and ground cable. Plug in the electrode lead to the jacks extending from the buffer chambers.

#### *Software set up*

Restart the Macintosh computer and open the ABI PRISM™ 377 collection. Choose genetic analysis program and create a run file. Set parameters as follows:

*Plate Check Module:* plate check A

*PreRun module:*

GS PR 36A-2400 with constant voltage of 1100V and gel temperature of 51 °C

*Run Module:*

GS Run 36A-2400 with constant voltage of 3000V and gel temperature of 51 °C

*Collection time:* 4h

*Gel's Matrix File:*

GS Set A for detection of fluorescent dye: FAM, JOE, TAMRA and ROX

*Well-to-Read distance:* 36 cm

*Lanes:* 50, 66 or 96 depending on the number of samples to be analysed

*Run Mode:* XL Scan

## Materials and methods

---

Before loading the gel, pre-run the gel to make sure the electrophoresis system is working properly and check for temperature equilibration. Pre-run the gel for 15 min under the conditions defined in the run file.

### *Sample sheet file*

Choose 'new' from the file menu, a sample sheet will appear by clicking the GeneScan® Run icon. Fill out the form with relevant information of samples and save it.

### *Loading the samples and starting the run*

Before loading the gel, click 'Pause' button to arrest the instrument. Pausing stops electrophoresis but maintains the temperature of the gel. 1.5 µl of each sample was loaded in an order identical to that of the sample sheet. After loading, click the 'Cancel' button to stop the pre-run and click the 'Run' button to start the run under the conditions defined in Run File.

### **Data collection and fragment sizing by ABI GeneScan<sup>®</sup>**

Automated AFLP fragment analysis was performed with GeneScan® analysis software version 2.1 (Perkin-Elmer/ABI). In automated genotyping on a 377 sequencer, the fluorescence labelled DNA fragments are separated on the polyacrylamide gel. At the bottom of the gel, a laser excites the fluorescent dyes when the fragments pass, and detectors collect the emission intensities at specific wavelengths as defined in GS Set A. The laser and detectors scan the bottom of the gel continuously during electrophoresis in order to build a gel image in which each lane has a colour pattern of peaks, each peak corresponding to a fragments of particular length. The ABI GeneScan® software, running on a Macintosh® computer, collects each fluorescent signal and then assigns a size to each fragment.

### *The analysis parameter*

*Analysis range:* full range

*Size call range:* all sizes

*Peak detection:* all peaks higher than 50 were taken as real fAFLP fragments.

### **GS-500 size standard**

Select a sample file containing GS-500 size standard and assign the known size in bp to each fragments of the size standard in the size standard dialog box.

### *Samples Analysis*

Open the 'project' window containing the data, select the size standard, the analysis parameters and colours for all samples and the internal size standard. Click the 'analyse'

button for automated sizing of fragments. Before proceeding to allele scoring, each sample file was checked to be sure that the sizes of the size standard correspond well with what has been defined.

### **Allele scoring by ABI GenoTyper<sup>TM</sup>**

In this study Genotyper<sup>TM</sup> analysis software version 2.0 (PE Applied Biosystems) was used for creation of data files. GenoTyper<sup>TM</sup> converts data from GeneScan<sup>®</sup> files into the Excel format required by downstream applications. The following steps are involved in the processes:

#### ***Importing result files from GeneScan<sup>®</sup> files***

#### ***Creating categories***

The categories define the peaks of interest that need to be labelled. In case of AFLP, where no information of the AFLP fragments was available, categories were defined which took all fragments ranging from 50 bp to 500 bp into account. One base pair corresponds, therefore, to one category with a tolerance of  $\pm 0.49$  bp.

#### ***Labelling peaks***

Peak labelling was done either quantitatively as peak height, or qualitatively as presence versus absence.

#### ***Making tables***

After auto-labelling, data tables for exporting were made which contained sample info, categories and other relevant information. These data can be exported in Excel format for QTL analysis.

## **2.14 Evaluation of phenotypic traits**

Phenotypic data were evaluated in field trials at two locations (Gransebieth and Windeby) over three years (1997, 1998, and 1999) by the potato breeding company SaKa-Ragis Pflanzenzucht GbR (Dr. Josef Strahwald). A short description is included here because the phenotypic data were an essential part of the QTL analysis.

Trait evaluation was carried out on 150-200 individuals of Z3 in trials 1997 and 1998. In trial 1999, evaluation was performed on 50 individuals, which were obtained by selecting 46 individuals with high and low tuber starch content in pairwise from each of 9 categories for foliage maturity (1-9 scale). Phenotypic trait evaluated varied in different trials, however, present study concentrates on the analysis for the following traits.

## Materials and methods

---

### *Starch content*

Tuber starch content in percent was determined by measuring specific gravity (weight in air – weight in water) using a starch balance (Meku, Wennigsen, Germany) based on the method of Lunden (1956).

### *Reducing sugars*

Reducing sugar content was measured by a Picric Acid Test. The degree of darkening of a solution of potassium picrate is a measure for the reducing sugar content of tubers. Picric acid reacts with reducing sugars to produce a red colour. In solutions containing excess picric acid, the red colour will combine with the yellow colour of the picric acid to produce a series of shades from yellow through orange to red. Reducing sugar content was estimated by comparing the colour with a series of colour standards which have been calibrated in terms of reducing sugar (see below).

CTC <sup>a</sup>	1	2	3	4	5	6	7	8	9
RSC <sup>b</sup>	0.059	0.12	0.20	0.28	0.38	0.49	0.67	0.95	1.35

<sup>a</sup> CTC=Colour Test Classification; <sup>b</sup> RSC=Reducing Sugar Content in %

### *Foliage Earliness (Maturity)*

Foliage earliness was scored using a scale of 1 to 9 (1=very early maturing, 9=very late maturing) at an appropriate growth stage of the plants.

## 2.15 QTL analysis

QTL analysis was carried out in collaboration with Dr. Ralf Schäfer-Pregl (MPIZ).

Several statistical methods can be used for the identification of QTLs depending upon the molecular marker types and scoring methods. In the present study, the statistical programs used for the QTL analysis include the t-test or one way ANOVA (analysis of variance) and GLM (general linear models). All analyses were performed with programs written using SAS software (SAS Institute Inc., 1990).

### *t-test*

A t-test was applied to the fAFLP data which were divided into two genotype classes such as presence and absence of a DNA fragment. The means of the two classes are tested for

significant differences by estimating the probability of no difference between the means compared.

### **ANOVA**

One-way ANOVA is an alternative to the t-test. This procedure tests the significance of the difference between the means of two subgroups by estimating the F value, the proportion of the variance between the two subgroups to the variance within each group, from which a *P* value or a LOD (logarithm of odds) value are derived. This procedure also gives the  $R^2$  value, indicating the proportion of the total variance that is explained at the marker locus considered.

### **GLM**

In case when multiple alleles are informative at given locus and can be distinguished in a segregating population, GLM was applied to the QTL analysis by examining the variance among the phenotypic trait means of multiple marker-genotype classes. Using this method, the amount of variance explained by allelic differences at a marker locus ( $R^2$ ) can also be calculated.

### ***Regression analysis***

Regression analysis was applied to quantitative fAFLP data. This analysis reveals the correlation between the peak height and the phenotypic values. The square of the correlation coefficient reflects the proportion of the variance explained by regression to the total variance.

### ***Non-parametric test***

For the category phenotypic data, the Mann-Whitney U-test and the Kruskal-Wallis test were employed for the association analysis. The former is the non-parametric equivalent to the t-test, while the latter one is the non-parametric counterpart of GLM.

### ***Thresholds for declaring a QTL present***

A marker-trait association was declared significant when the mean or median values of the two marker classes were significantly different at  $P < 0.05$  in two locations or two independent tests. Otherwise  $P < 0.01$  was selected as threshold.



### 3. RESULTS

#### 3.1 Mapping functional genes involved in carbohydrate metabolism and transport

##### 3.1.1 PCR amplification of functional genes through an STS approach

Genes involved in starch and sucrose metabolism, metabolite transport, Calvin cycle, oxidative pentose phosphate pathway, glycolysis and TCA cycles were explored in public databases of EMBL and GenBank for their sequences in potato, tomato or other species. 51 gene-specific primer pairs were designed (Table 3-1, page 51-55), of which 45 primer pairs were derived directly from the DNA sequence of potato genes and three pairs from tomato genes (*Eno*, *Ndpk*, and *Glo*). For genes *Ant*, *Rca* and *Pgk<sub>cp</sub>*, which were not available either as a potato or as a tomato sequence, degenerate primers were designed based on sequence motifs conserved between the homologous genes of three to five other plant species (Schneider *et al.* 1999, Van Campenhout *et al.* 2000). The consensus sequence of a gene can be revealed by sequence analysis using the *Pile-Up* program of *GCG* software package (University of Wisconsin Genetics Computer Group, version 8). The example of a multiple sequence pileup of sequences of the *Pgk* gene from several species is shown in Figure 3-1 (on page 47-49). Using potato genomic DNA as a template, most primer pairs generated a single fragment, only a few produced two fragments. Most frequently the PCR products were larger than expected from the transcript sequence, indicating the presence of introns.

	551					600
u37701	CGAAAGCTGA	TGATTGTATT	GGCCCAGAAG	TGGAAAGCTT	GGTGGCTTCT	
z48977	TGAAGGTTGA	GGACTGCATT	GGTCCAGAAG	TTGAGAAGTT	GGTTGCTTCA	
x68430	TGAAGGCTGA	TGATTGCATT	GGTCCAGATG	TTGAGAAGTT	GGTTGCTGAA	
x15233	AAAAAGCAGA	AGATGTTATC	GGCCCAGAAG	TTGAGAAATT	GGTGGCTGAC	
	** *	** *	** *	** *	** *	** *
	601					650
u37701	CTACCTGAAG	GTGGAGTTTT	GCTTCTTGAG	AACGTCAGGT	TTTACAAGGA	
z48977	CTTCCCGAGG	GTGGTGTTC	TCTTCTCGAG	AACGTGAGAT	TCTACAAGGA	
x68430	CTCCCGAAG	GTGGCGTTC	TCTCCTTGAG	AATGTTAGGT	TCTACAAGGA	
x15233	CTGGCAAATG	GTGCTGTTTT	GCTCCTGGAA	AACGTAAGAT	TTTACAAGGA	
	** *	** *	** *	** *	** *	** *
	651					700
u37701	GGAAGAGAAG	AACGATCCTG	AGTTTGCTAA	GAAGCTTGCT	TCTCTAGCTG	
z48977	GGAAGAGAAG	AACGAACCTG	AGTTTGCAAA	GAACTTGCA	TCATTGGCAG	
x68430	AGAGGAAAAG	AATGATCCAG	AATTTGCAAA	GAAGCTTGCT	TCCTTGGCAG	
x15233	AGAGGAGAAG	AATGACCCAG	AGTTTGCAAA	GAAGCTTGCG	TCACTGGCAG	
	** *	** *	** *	** *	** *	** *

## Results

```

701
u37701 ACCTTTATGT CAATGATGCT TTCGGAAGCTG CTCACAGAGC TCATGCTTCT
z48977 ATCTTTATGT GAATGATGCA TTCGGTACAG CTCACAGAGC ACATGCCTCT
x68430 ATCTGTATGT GAATGATGCT TTTGGCACTG CCCACAGAGC TCATGCATCA
x15233 ATCTGTTTGT AAACGATGCA TTCGGAACAG CCCACAGAGC ACATGCCTCG
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
751
u37701 ACCGAAGGAG TCACTAAGTT CTTGAAGCCT TCAGTTGCTG GTTTCCTTTT
z48977 ACAGAGGGAG TTAATAAATT TTTGAAGCCT TCTGTTGCAG GTTTCCTCTT
x68430 ACCGAGGGAG TTACCAAATT TTTGAAGCCA TCTGTTGCTG GCTTCCTCTT
x15233 ACTGAGGGAG TTACCAAGTT CTTGAAGCCT TCTGTTGCTG GATTTCCTTTT
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
801
u37701 GCAAAAGGAA CTGGACTACC TAGTTGGTGC TGTTTCAAAC CCAAAGAGAC
z48977 ACAAAGGAA TTGGACTATT TAGTCGGGGC GGTTTCAAAT CCAAAGAGGC
x68430 ACAGAAGGAA TTGGACTACC TTGTTGGAGC TGTTTCTAAC CCAAAGAGAC
x15233 GCAGAAGGAA CTTGACTACC TGGATGGAGC TGTTTCAAAC CCTAAGCGCC
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
degenerate primer
851
u37701 CATTTGCAGC CATAGTGGGT GGTTCCAAAG TCTCATCCAA GATTGGAGTT
z48977 CATTTGCTGC TATTGTGGGT GGTTCAAAGG TTTTCATCCAA GATTGGAGTG
x68430 CATTTGCTGC TATAGTTGGT GGTTCAAAGG TCTCTTCCAA AATTGGAGTA
x15233 CATTTGCTGC CATTTGTGGGT GGCTCAAAGG TGTCATCCAA GATTGGGGTT
CATTTGC * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
901
u37701 ATTGAATCGC TTCTGGAGAA GTGTGATATT CTTCTTCTTG GTGGTGGAAT
z48977 ATCGAATCAC TTTTAGAGAA ATGTGATATA TTGCTTTTGG GTGGAGGAAT
x68430 ATTGAGTCAC TTTTGGAGAA ATGTGATATC CTACTCCTTG GAGGAGGAAT
x15233 ATCGAATCCC TGTGGAGAA GTGTGACATC CTTCTTTTGG GTGGTGATAT
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
951
u37701 GATCTTCACA TTCTACAAGG CACAGGGTCT TTCAGTTGGT TCGTCCCTTG
z48977 GATCTTTACC TTCTACAAGG CTCAGGGTCT TTCAGTTGGT TCCTCCTTGG
x68430 GATCTTTACA TTCTACAAG CTCAAGGTAT GTCAGTTGGA TCTTCTCTGG
x15233 GATCTTCACA TTTTACAAGG CACAAGGACT CTCAGTTGGT TCTTCCTTGG
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
1001
u37701 TTGAAGAAGA CAAGCTTGAA TTGGCTACAG AACTCCTTGC CAAAGCTAAG
z48977 TTGAGGAAGA CAAACTAGAA CTCGCTACAT CACTCCTAGA GAAGGCCAAG
x68430 TTGAAGAGGA TAAGCTTGAC CTTGCAACTT CTCTCCTCGC CAAGGCTAAG
x15233 TTGAGGAAGA TAACTTGAG CTGGCAACAT CTCTCCTTGC AAAGGCCAAG
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
1051
u37701 GCCAAAGGAG TCTCTCTTTT GTTGCCAACA GATGTTGTAG TTGCTGACAA
z48977 GCGAAAGGAG TCAGTCTCTT GTTACCATCT GATGTTGTGA TTGCAGATAA
x68430 GAAAAGGGCG TGTCTCTGTT GTTGCCCTACT GATGTCGTTA TTGCGGACAA
x15233 GCAAAGGGTG TCTCCCTTCT GTTGCCATCT GACGTTATCA TTGCTGATAA
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
1101
u37701 GTTTGCTCCT GATGCCAACA GCAAGATTGT GCCTGCATCA GGCATTGAGG
z48977 ATTTGCTCCT GATGCAAACA GCAAGATTGT GCCTGCATCT GCTATCCCAG
x68430 GTTTGCTGCT GATGCAGACA GCAAGATTGT CCCC GCATCT GGAATCCCTG
x15233 GTTTGCTCCT GATGCTAACA GTCAGACCGT CCCTGCATCT GCAATTCTCTG
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
1151
u37701 ACGGATGGAT GGGACTGGAC ATTGGTCCAG ACTCTATCAA AACTTTCAAC
z48977 ACGGTTGGAT GGGGTTGGAC ATTGGACCAG ACTCTGTAA GACTTTCAAC
x68430 ACGGATGGAT GGGATTTGGAC ATTGGACCTG ATTCTATTAA GACATTTAGT
x15233 ATGGTTGGAT GGGGCTGGAC ATTGGCCCAG ATTCAGTCAA AACATTTAAT
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

## Results

	1201				1250
u37701	GAAGCTCTGG	ACACAACACA	AACAGTCATT	TGGAATGGAC	CTATGGGGAGT
z48977	GATGCCTTGG	ATACCACAAA	AACAGTGATC	TGGAATGGAC	CTATGGGGGT
x68430	GAAGCTTTGG	ACACTACCCA	GACCGTCATC	TGGAATGGTC	CCATGGGGAGT
x15233	GATGCCCTGG	ACACAACACA	GACAATCATT	TGGAACGGAC	CAATGGGTGT
	** ** *	*** * * * *	** * *	***** * *	* ***** *
	1251				1300
u37701	TTTCGAGATG	GAAAAGTTTG	CGGCTGGAAC	AGAGGCGATA	GCGAATAAAC
z48977	GTTTGAATTT	GACAAGTTTG	CTGTTGGAAC	AGAGGCAATT	GCAAAGAAGC
x68430	TTTTGAATTC	GAAAAGTTTG	CTGCTGGTAC	AGAGGCTATT	GCTAAGAAGC
x15233	CTTTGAATTT	GACAAGTTTG	CAGTAGGAAC	TGAGTCTATT	GCAAAGAAGT
	** * *	** *****	* * ** *	*** * *	** ** *
	1301				1350
u37701	TAGCAGAGCT	AAGTGAAAAA	GGAGTGACAA	CGATAATAGG	AGGAGGAGAC
z48977	TCGCGGACTT	AAGTGGGAAA	GGAGTGACGA	CTATCATTGG	AGGTGGAGAT
x68430	TAGAGGAGAT	AAGCAAGAAG	GGTGCAACCA	CAATTATCGG	TGGTGGTGAC
x15233	TGGCCGAGCT	TAGCAAAAAG	GGTGTGACAA	CTATCATTGG	AGGCGGAGAC
	* * *	* **	** * ** *	* ** ** *	** ** ** *
	1351				1400
u37701	TCAGTGGCTG	CAGTGGAGAA	AGTAGGAGTA	GCAGGAGTCA	TGAGTCACAT
z48977	TCTGTTGCAG	CTGTTGAGAA	AGTTGGAGTT	GCTAGCGTGA	TGAGCCACAT
x68430	TCAGTTGCAG	CAGTAGAGAA	GGTAGGAGTG	GCAGAGGCAA	TGAGCCACAT
x15233	TCCGTTGCGG	CTGTTGAGAA	GGTGGGAGTG	GCTGATGTTA	TGAGCCACAT
	** ** ** *	* ** *****	** *****	** * *	**** ** *
	1401				1450
u37701	CTCCACTGGT	GGTGGAGCCA	GCTTGGAGCT	GTTGGAAGGA	AAAGTACTTC
z48977	ATCCACTGGT	GGTGGTGCCA	GTTTGGAGCT	ACTGGAAGGC	AAGGTGCTCC
x68430	ATCAACTGGT	GGTGGAGCTA	GTTTGGAGTT	GTTAGAAGGG	AAACAGCTTC
x15233	CTCA <b>ACTGGT</b>	<b>GGTGGCGCCA</b>	<b>GCTTGGAGTT</b>	GTTGGAAGGA	AAGGAGCTTC
	** <b>TGACCA</b>	<b>CCACDCGRT</b>	<b>CRAA</b> **** *	* *****	** ** *
	degenerate primer				
	1451				1500
u37701	CCGGTGTGAT	CGCCCTTGAT	GAAGCAATCC	CAGTCACTGT	TTAGAAACTC
z48977	CCGGTGTGAT	TGCTCTTGAT	GAAGCAGACG	CCCCCGTTGC	TGTGTAAAC
x68430	CTGGAGTACT	TGCTCTTAAT	GAGGCCGACC	CAGTTCCGGT	TTAAGACCAC
x15233	CTGGAGTCGT	TGCACTTGAT	GAAG. .GTGT	CATGACGAGG	TCCGTGACCG
	* ** ** *	** ** ** *	** *	*	*
	1501				1550
u37701	TTACCATCTA	TAAGGCAAAC	GCTGTATACT	TAAAAGAGTT	TCTTTCGCTT
z48977	AATTTGTACT	AATTCTTTTT	TCTCCGGGGT	CAACATAATC	AGTGGAATTT
x68430	TCTATACAAA	TTCTTTTTTT	CCCCCTCTCT	TCACAGGTTT	ATCCTCTTAA
x15233	TATGAGGCTA	AGCTTCATTT	GTTGCATCTT	AATTCCTTTC	ATGTACCGTT
			*	*	

**Figure 3-1 Multiple alignment of partial cDNA sequences of the gene coding for phosphoglycerate kinase (PGK) from 4 different species showing the conserved sequence motifs used as primers for amplifying homologous genes in potato.** U37701, Z48977, X68430 and X15233 are cDNA sequences from *Arabidopsis*, tobacco, spinach and wheat, respectively. Asterisks indicate identical positions. Sequence motifs adopted for primers are highlighted in bold.

### 3.1.2 Polymorphism Identification

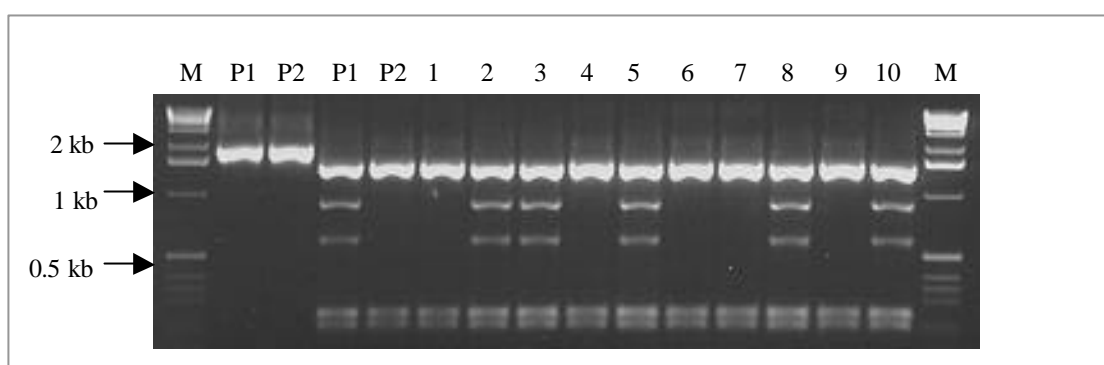
Most primer pairs were selected to flank introns which are potentially more variable when compared to exons (Tragoorung *et al.* 1992; Konieczny and Ausubel 1993; Corte-Real *et al.* 1994). This allowed efficient identification of polymorphisms by SCAR or CAPS. In total, four primer pairs for genes *SssI*, *Pk*, *Me* and *Ndpk* generated directly polymorphic PCR

## Results

products of which *Me* detected two loci (SCAR markers, Table 3-1, page 51-55). PCR products of 45 genes were digested with a 4-base pair cutter restriction enzyme to generate segregating DNA fragments (CAPS markers, Figure 3-2). 43 of the 45 genes were mapped as CAPS markers using one of the following enzymes *TaqI*, *RsaI* and *AluI*. The remaining two genes were mapped using *HaeIII* and *DdeI*. The most informative restriction enzyme was *TaqI* followed by *RsaI* and *AluI* (Table 3-1, page 51-55). PCR products of all the genes were restricted with at least these three enzymes for CAPS to maximize the identification of loci even for the ones showing directly SCAR polymorphism. However, the increase of polymorphic fragments produced by different restriction enzymes did not increase the number of loci identified. Frequently, these polymorphic fragments, which were derived from the same genes, tagged an alternative allele of one parent or the other or common alleles of both. In the case of *Pain-1* and *Eno* for example, CAPS assays with the enzymes *DdeI* and *TaqI* for *Pain-1* and *AluI* and *TaqI* for *Eno* resulted in 3 polymorphic fragments with different segregation patterns which were all mapped at a single locus on linkage groups (LG) III and IX, respectively. The only exceptions was *Ppe* which identified three loci, two (LG III, LG V) by *AluI* and one (LG XII) by *TaqI*. Other genes with two loci being identified are *Pha1*, *Glo* and *Fum1* by *TaqI* and *RsaI* (Table 3-1, on page 51-55).

The PCR product of only one gene (*Tk*), which may contain no introns in the region investigated, did not segregate either as SCAR or as CAPS marker. Alternatively, by RFLP analysis using the PCR product as the probe, the *Tk* gene was successfully mapped on LG V (Table 3-1, on page 51-55).

(continue to page 56)



**Figure 3-2 CAPS marker for the  $\alpha$ -glucan phosphorylase gene *Stp23* which was mapped on potato chromosome III.** The uniform 2 kb PCR products, shown for the parents P1 and P2 of the LH mapping population (2<sup>nd</sup> and 3<sup>rd</sup> lane from left), were polymorphic after digestion with *TaqI* (4<sup>th</sup> and 5<sup>th</sup> lane) and segregated in the F1 plants (6<sup>th</sup> to 15<sup>th</sup> lane from left). M = size marker.

## Results

**Table 3-1 Genes placed onto the potato molecular function map**

Enzyme	EC number	Accession number	Reference	Gene	Linkage group	Primer, 5' to 3'	Population <sup>a</sup>	PCR product kb	Marker Assay <sup>b</sup>
<b>A: Starch synthesis and degradation (chloroplasts, amyloplasts)</b>									
ADP-glucose pyrophosphorylase B <sup>e</sup>	2.7.7.27	X55155	Mueller-Rober <i>et al.</i> 1990	<i>AGPaseB</i>	VII, XII	-	1,2,3,4	-	R, TaqI
ADP-glucose pyrophosphorylase S <sup>e</sup>	2.7.7.27	X61186	Nakata <i>et al.</i> 1991	<i>AGPaseS</i>	I, IV, VIII	-	1,2,3,4	-	R, RsaI, TaqI
Granule bound starch synthase I <sup>e</sup>	2.4.1.21	X52416	Rohde <i>et al.</i> 1990	<i>GbssI (wx)</i>	VIII	-	1, 3, 4	-	R, RsaI
Soluble starch synthase I	2.4.1.21	Y10416	Abel <i>et al.</i> unpublished	<i>SssI</i>	III	f-cttcttactgcagacctggaac r-ctgtcagtatccgatcagcaac	3	2.4 2.6	S
Granule bound starch synthase II	2.4.1.21	X87988	Edwards <i>et al.</i> 1995	<i>GbssII</i>	II	f-gctgcaagtgcgtgatgaatcga r-ttagaccatggagcgcatctctg	3	1.8	C, TaqI
Soluble starch synthase III	2.4.1.21	X94400	Abel <i>et al.</i> 1996	<i>SssIII</i>	II	f-aacaaaagtgcagtcctctctc r-aaatcccaccatctctctctc	3	1.3	C, AluI
Starch branching enzyme I <sup>e</sup>	2.4.1.18	Y08786	Khoshnoodi <i>et al.</i> 1996	<i>SbeI</i>	IV	-	1, 3, 4	-	R, TaqI
Starch branching enzyme II	2.4.1.18	AJ000004	Larsson <i>et al.</i> 1998	<i>SbeII</i>	IX	f-ctcgtagtgtctacaggtatcac r-tgatggagtggtccatacgtg	3	1.8	C, AluI
Disproportionating enzyme	2.4.1.25	X68664	Takaha <i>et al.</i> 1993	<i>Dpe-P</i>	IV	f-cactacttttcaatctcctatccc r-gcatagtcacgaactttttcc	3	3.0	C, TaqI
Starch-granule-bound protein R1	-	Y09533	Lorberth unpublished	<i>Gb-R1</i>	V	f-tccatcctgagactggagatac r-acttgtagtgcaggactggaag	4	1.7	C, RsaI
Debranching enzyme	3.2.1.41	A52190	Kossmann patent	<i>Dbe</i>	XI	f-ctgatgtcagcatctatgagct r-gatacgacaaggaccatttgca	3 4	2.1	C, TaqI R, AluI
$\alpha$ -Amylase <sup>e, g</sup>	3.2.1.1	-		<i>AmyZ</i>	IV	-	1, 2	-	R, TaqI
$\alpha$ -Glucosidase	3.2.1.20	AJ001374	Taylor <i>et al.</i> 1998	<i>Agl</i>	IV	f-accaagctgtgtgtaaccagag r-gcagttgcgaataactgtggca	3	0.8	C, AluI
H-type starch phosphorylase	2.4.1.1	- <sup>f</sup>		<i>StpH</i>	IX	f-gcatactatgctgctactgctg r-gcacatcatatgcaagagcctg	3	1.8	C, RsaI
L-type starch phosphorylase	2.4.1.1	X73684	Sonnenwald <i>et al.</i> 1995	<i>StpI</i>	V	f-acacactatgttctgcttctctc r-actatcctccacctcaaccttc	4	0.8	C, TaqI RsaI

## Results

**Table 3-1 Genes placed onto the potato molecular function map (continued)**

Enzyme	EC number	Accession number	Reference	Gene	Linkage group	Primer, 5' to 3'	Population <sup>a</sup>	PCR product kb	Marker Assay <sup>b</sup>
$\alpha$ -Glucan phosphorylase	2.4.1.1	D00520	Nakano <i>et al.</i> 1989	<i>Stp23</i>	III	f-atggcgactgcaaatggagca r-ccatactgtacctaagtccatag	4	1.6	C, TaqI, RsaI, AluI
<b>B: Sucrose metabolism (cytosol, apoplast)</b>									
UDP-glucose pyrophosphorylase <sup>c</sup>	2.7.7.09	U20345	Borovkov unpublished	<i>UGPase</i>	XI	-	1, 2, 4	-	R, TaqI
Fructose-1,6-bisphosphatase	3.1.3.11	X76946	Zrenner <i>et al.</i> 1996	<i>Fbp<sub>cy</sub></i>	IV	f-tgcaggaggagaagatcaaaagaaac r-tgaagaacctcagcgggatac	3	1.4	C, RsaI TaqI
Pyrophosphate fructose-6-phosphate 1-phosphotransferase, $\alpha$ subunit	2.7.1.90	M55190	Carlisle <i>et al.</i> 1990	<i>Pfp-a</i>	IV	f-cggacaaaagatcagattagaacgac r-ccctcaggcaacaggataacac	3	2.5	C, AluI
Pyrophosphate fructose-6-phosphate 1-phosphotransferase, $\beta$ subunit	2.7.1.90	M55191	Carlisle <i>et al.</i> 1990	<i>Pfp-b</i>	II	f-gtcatgatagatgctcgatcaac r-acatcagcaatgtagtcgtaac	4	2.5	C, RsaI
Sucrose phosphate synthase <sup>g</sup>	2.4.1.14	X73477	Zrenner <i>et al.</i> 1995	<i>Sps</i>	VII	f-gaaagagggtcgcagagaagcag r-caagtgtggtgaccagtgaag	2, 4	1.3	C, TaqI R, TaqI
Sucrose synthase 3	2.4.1.13	U24088	Fu <i>et al.</i> 1995	<i>Sus3</i>	VII	f-catgacaaggaaagcatgacccc r-gcaaagtaaatcttatacatgtagc	4	1.2	C, TaqI
Sucrose synthase 4 <sup>k</sup>	2.4.1.13	U24087	Fu <i>et al.</i> 1995	<i>Sus4</i>	XII	f-caagctgacctggacaccacagt r-accacattgaaaaccataggaattct	3	1.3	C, RsaI
Soluble acid invertase	3.2.1.26	X70368	Zrenner <i>et al.</i> 1996	<i>Pain-1</i>	III	f-gtcttcgataagacttttcgag r-cagtgtgtcgggtctctaaagt	3	2.0	C, DdeI TaqI
Apoplastic invertase	3.2.1.26	Z22645	Hedley <i>et al.</i> 1994	<i>Inv<sub>ap</sub></i>	IX, X	-	1	-	R, TaqI
<b>C: Transport (membranes)</b>									
Sucrose transporter	-	X69165	Riesmeier <i>et al.</i> 1993	<i>Sut1</i>	XI	f-agcttccatagctgctggtgtt r-cggactaagggttaaggctatgg	1, 3	1.2	C, RsaI R, TaqI
Triose phosphate translocator <sup>e</sup>	-	-		<i>Tpt</i>	X	-	2	-	R, RsaI
Glucose-6-phosphate translocator	-	AF020816	Kammerer <i>et al.</i> 1998	<i>Gpt</i>	V	f-ggctcacacaattggtcatgtg r-ccaagattgcaatagcagcacc	4	1.4	C, HaeIII
Plasma membrane H <sup>+</sup> -ATPase 1 <sup>e</sup>	3.6.1.34	X76536	Harms <i>et al.</i> 1994	<i>Phal</i>	III, VI	f-tcctggagatggtgtctactct r-gcagtatcaatggcatcctggt	2, 3	0.8	C, TaqI R, TaqI

## Results

**Table 3-1 Genes placed onto the potato molecular function map (continued)**

Enzyme	EC number	Accession number	Reference	Gene	Linkage group	Primer, 5' to 3'	Population <sup>a</sup>	PCR product kb	Marker Assay <sup>b</sup>
Plasma membrane H <sup>+</sup> -ATPase 2 <sup>e</sup>	3.6.1.34	X76535	Harms <i>et al.</i> 1994	<i>Pha2</i>	VII	f-caaacatgtaccgcgtcagcatc r-agctatcaggcattggagatgg	1, 4	0.9	R, TaqI C, TaqI
Adenylate transporter	-	X65549 X62123 D12637 X95864 X68592		<i>Ant</i>	XI	f-tggagaggaaacactgcYaatgt <sup>1</sup> r-atgttRgcaccagcWcccttga <sup>1</sup>	4	1.1 + 0.9	C, TaqI
Inorganic phosphate transporter 1	-	X98890	Leggewie <i>et al.</i> 1997	<i>Pt1</i>	IX	f-tcagcagctgttaatggagtcg r-tgcaacaccttggacatgctgt	4	0.6	C, RsaI
Inorganic phosphate transporter 2	-	X98891		<i>Pt2</i>	III	f-cgccacgatcatgtctgaatac r-taaacgccacgggtgaaccagta	4	0.7	C, AluI
<b>D: Calvin cycle, photorespiration (chloroplasts)</b>									
Ribulose biphosphate carboxylase activase	-	J03610 AF037361 X14212 Z21794		<i>Rca</i>	X	f-acacYgtMaacaaccagatg <sup>1</sup> r-actctcttgacattctcttgc	4	0.7	C, TaqI
Ribulose biphosphate carboxylase small subunit <sup>e</sup>	4.1.1.39	- <sup>1</sup>		<i>rbcS</i>	II (2) III	-	1,3,4	-	R, TaqI, RsaI
Phosphoglycerate kinase	2.7.2.3	U37701 Z48977 X68430 X15233		<i>Pgk<sub>cp</sub></i>	VII	f-aaYccWaagMgVccatttgc <sup>1</sup> r-aaRctRgcDccaccaccagt <sup>1</sup>	4	0.9	C, TaqI
Fructose-1,6-bisphosphatase	3.1.3.11	- <sup>h</sup>	Kossmann <i>et al.</i> 1992	<i>Fbp<sub>cp</sub></i>	IX	f-tactggagttcaaggtgctgtc r-tgcagagtaaggcttccgacta	3	1.6	C, TaqI
Transaldolase	2.2.1.2	U95923	Moehs <i>et al.</i> 1996	<i>Tall</i>	XI	f-attccttgtgtgcaaatgctcc r-cgactaacgaagaatgaagcaac	4	1.5	C, TaqI
Transketolase	2.2.1.1	Z50099	Teige <i>et al.</i> 1995	<i>Tk</i>	V	f-gagacacggaaattgctttcac r-tcagcactaccaccaaggaaac	4	0.6	R, TaqI
Pentose-5-phosphate 3-epimerase	5.1.3.1	Z50098	Teige <i>et al.</i> 1995	<i>Ppe</i>	III, V, XII	f-tccgtccatcttttctgctaac r-aactccaccatcaacttcaatc	3	3.0	C, AluI TaqI
Glycolate oxidase	1.1.3.15	X92888	Speirs <i>et al.</i> 1995	<i>Glo</i>	VII, X	f-cagcgttgaggaggttgcttca r-gacagccactcaatgccatagt	4	1.7	C, TaqI

## Results

**Table 3-1 Genes placed onto the potato molecular function map (continued)**

Enzyme	EC number	Accession number	Reference	Gene	Linkage group	Primer, 5' to 3'	Population <sup>a</sup>	PCR product kb	Marker Assay <sup>b</sup>
<b>E: Glycolysis, oxidative pentose phosphate pathway (cytosol)</b>									
Fructokinase	2.7.1.4	Z12823	Smith <i>et al.</i> 1993	<i>Fk</i>	VI	f-gctttggcggttcgtgactctac r-agtgggtgcaacagtcttcacg	3	2.0	C, TaqI
Hexokinase <sup>g</sup>	2.7.1.1	X94302	Clericus unpublished	<i>Hxk</i>	III	f-gattatattgcggctgagcttgc r-catctgccgttgacagagtatg	2, 3	0.8	C, RsaI R, TaqI
Glyceraldehyde 3-phosphate dehydrogenase <sup>d</sup>	1.2.1.12	U17005	Schneider <i>et al.</i> 1995	<i>Gap C</i>	V	-	1	-	R, TaqI
Enolase	4.2.1.11	X58108	Van der Straeten <i>et al.</i> 1991	<i>Eno</i>	IX	f-cttggtgcaaatgccatccttg r-cagcttcaatactctcggtcac	3	2.1 + 1.9	C, AluI TaqI
Pyruvate kinase	2.7.1.40	X53688	Blakeley <i>et al.</i> 1990	<i>Pk</i>	IV	f-tcactgtatccacagactataacc r-cactctccccacttaacataac	3	0.9 0.8	S
Lactate dehydrogenase 2 <sup>g</sup>	1.1.1.28	AF067859	Dunford unpublished	<i>Ldh2</i>	VIII (3)	-	2	-	R, TaqI
Phosphoenolpyruvate carboxylase	4.1.1.31	X67053	Merkelbach <i>et al.</i> 1993	<i>Ppc</i>	X, XII	-	1	-	R, TaqI
Glucose-6-phosphate dehydrogenase	1.1.1.49	X74421	Graeye <i>et al.</i> 1994	<i>G6pdh</i>	II	f-tcttctattttgtcttctctcc r-actcgtatttctgcctttcttg	3	1.6	C, DdeI
<b>F: TCA cycle (mitochondria)</b>									
Pyruvate dehydrogenase, E1 $\alpha$	1.2.4.1	Z26949	Grof <i>et al.</i> 1995	Pdh-E1 $\alpha$	V	f-tcaacaagccgagccattaac r-gattagcagcaccatcaccatac	3	0.6	C, RsaI
Citrate synthase	4.1.3.7	X75082	Landschutze <i>et al.</i> 1995	<i>Cis</i>	I	f-ggtcacttggttgetagtgtt r-ctttcgcagaactccatgtcca	3	1.5	C, TaqI
Aconitase	4.2.1.3	X97012	Surpili unpublished	<i>Aco</i>	VII	f-catggctcctgaatatggtgca r-atacaagtgggtgcagccatagc	4	1.3	C, AluI HpaII
Isocitrate dehydrogenase 1	1.1.1.42	X75638	Fieuw unpublished	<i>Icdh-1</i>	I	f-aaaatatccccagacttgtccc r-aacagcgtcaatgaactcttcc	3	1.8	C, AluI
Fumarase	4.2.1.2	X91615	Nast unpublished	<i>Fum1</i>	VII, IX	f-catgtgcatccaaatgacctgt r-gaagcagccagggtatttaagg	4	1.9	C, RsaI
Malate dehydrogenase	1.1.1.39	Z23023	Winning <i>et al.</i> 1994	<i>Mdh</i>	VIII	f-acagtggattgttcagacgtcc r-gtcaagccgatgatcttgcagt	4	3.0	C, AluI



## Results

**Table 3-1 Genes placed onto the potato molecular function map (continued)**

Enzyme	EC number	Accession number		Gene	Linkage group	Primer, 5' to 3'	Population <sup>a</sup>	PCR product kb	Marker Assay <sup>b</sup>
NAD <sup>+</sup> -dependent malic enzyme	1.1.1.39	Z23002	Winning <i>et al.</i> 1994	<i>Me</i>	III, V	f-ccttggttcaataaggacacgg r-caacgagcatgaacagcttcca	4	1.1 1.2	S
<b>G: Miscellaneous</b>									
Soluble inorganic pyrophosphatase <sup>°</sup>	3.6.1.1	Z36894	Du Jardin <i>et al.</i> 1995	<i>Ppal</i>	VIII, XII	-	1	-	R
GDP-mannose pyrophosphorylase	2.7.7.13	AF022716	Keller unpublished	<i>GMPase</i>	III	f-gaaactgaaccacttggcactg r-ccctctccaatcttggcagatt	4	0.6	C, AluI
Uncoupling protein	-	Y11220	Laloi <i>et al.</i> 1997	<i>Ucp</i>	IX	f-gtategtacctgggttacatcg r-acattccaagatcccaagcgtc	3	2.5	C, TaqI
Pectin methyl esterase	3.1.1.11	S66607	Pear <i>et al.</i> 1993	<i>Pme</i>	VII	f-acttgctgaagcagttgctgca r-ttagccacagtgaacggcatag	4	1.5, 1.2, 1.0	C, TaqI
3-deoxy -D-arabino-heptulosonate 7-phosphate synthase	4.1.2.15	M95201	Zhao <i>et al.</i> 1993	<i>ShkB</i>	XI	f-catcttctccataaccctttacc r-gctcacagttctccaaaaatc	3	1.2	C, RsaI
Nucleoside diphosphate kinase	2.7.4.6	X75324	Harris <i>et al.</i> 1994	<i>Ndpk</i>	IX	f-atgatcaagcctgatggtgtcca r-aagagtgaaggctgctctgcca	3	1.1 0.4	S
β-1,3-glucanase, basic <sup>°</sup>	3.2.1.39	AJ009932	Talarczyk <i>et al.</i> 1998	<i>Glu B</i>	I	f-cagcctatcggagtatgctatg r-cattgaatccacaagggcacgcg	3 1	0.7	C, RsaI R
β-1,3-glucanase, acidic	3.2.1.39	-		<i>Glu A</i>	I	-	1	-	R, TaqI
Chitinase, class I, basic <sup>°</sup>	3.2.1.14	-		<i>Cht B</i>	X	-	1, 3, 4	-	R, TaqI
Protein kinase	2.7.1.37	X90990	Meyer <i>et al.</i> 1996	<i>Prk</i>	II	f-ccatctgagtcaggaaatcacct r-acattgatcgagcactccttgg	3	1.6	C, RsaI

<sup>a</sup> Four populations were used for mapping: 1 = BC916<sup>2</sup>, 2 = F1840 (Gebhardt *et al.* 1991, 1994), 3 = K31, 4 = LH (Schäfer-Pregl *et al.* 1998)

<sup>b</sup> C = CAPS, S = SCAR, R = RFLP

<sup>c</sup> The sequence of RFLP marker *CP58* originating from leaf cDNA (Gebhardt *et al.* 1991) codes for the 3' end of potato UGPase.

<sup>d</sup> The sequence of the cold regulated cDNA *CI13* isolated from cold stored tubers (van Berkel *et al.* 1994) codes for GAP C (Schneider 1995).

<sup>e</sup> Map position of this gene has been reported previously (Gebhardt *et al.* 1994, Du Jardin *et al.* 1995), they have been included in the Table based on their role in carbohydrate metabolism or transport.

<sup>f</sup> Sequence from Mori *et al.* 1991

<sup>g</sup> Potato EST clone (unpublished results from this laboratory) was used as probe for RFLP mapping

<sup>h</sup> Sequence from Kossmann *et al.* 1992

<sup>i</sup> Sequence from Eckes *et al.* 1985, and Wolter *et al.* 1988

<sup>k</sup> The *Sus4* gene of potato corresponds to *sh1* of maize (Gebhardt *et al.* 1991, Werr *et al.* 1985)

<sup>l</sup> Y = C + T, W = A + T, M = A + C, R = A + G, V = A + C + G, D = A + G + T

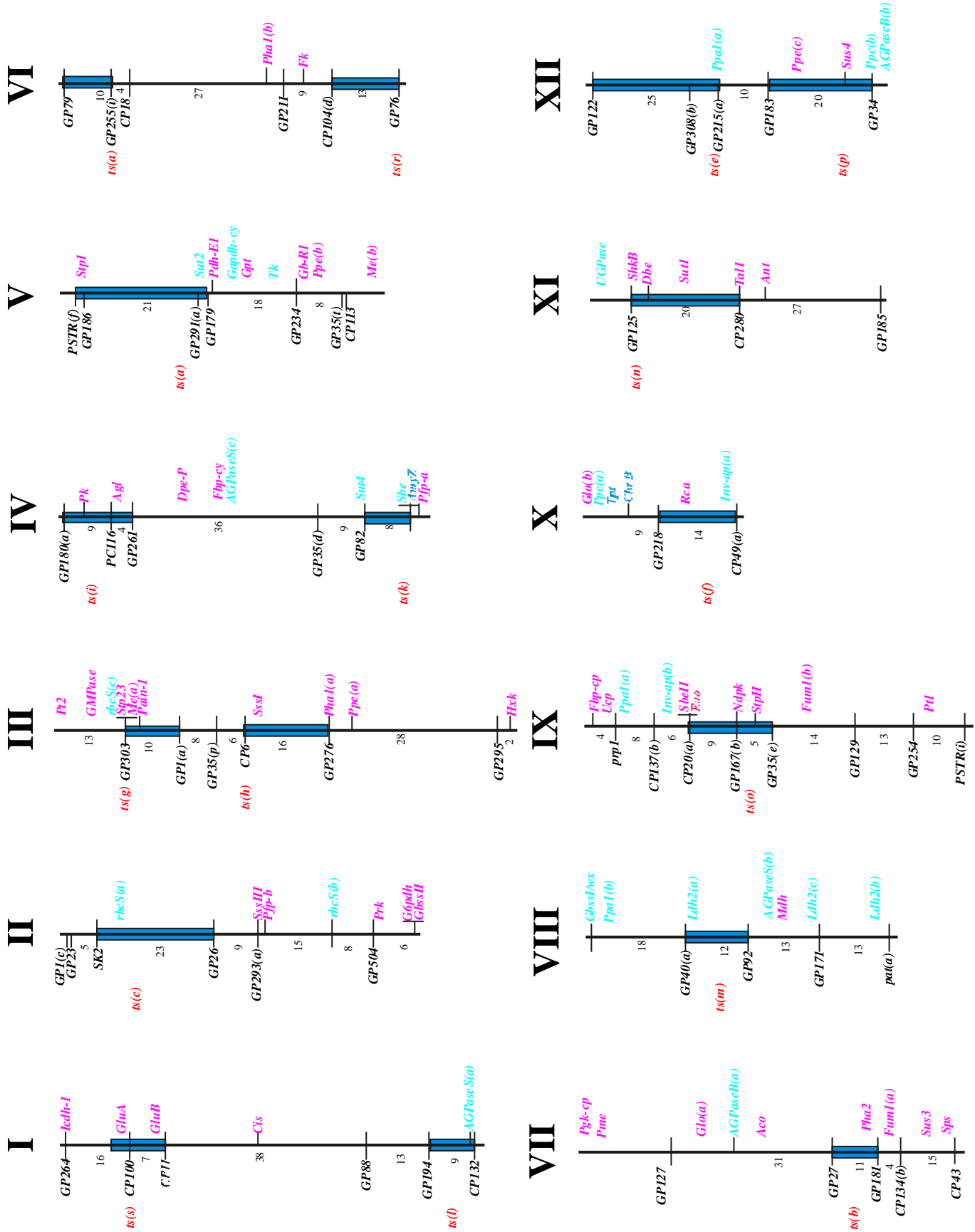
(continued from page 50)

### 3.1.3 Segregation analysis and function map construction

To localise the genes, segregation data were created by scoring 1 for the presence and 0 for the absence of DNA fragments. When multiple polymorphic fragments were detected for a single gene, they were scored separately and all were subjected to linkage analysis without considering their possible allelism. The software program MAPRF<sup>®</sup> (E. Ritter, Version 2.1) was employed for segregation and linkage analysis in conjunction with previously generated RFLP marker framework data sets of the same mapping populations K31 or LH. Depending on the parental genotypes, two segregation patterns were expected for any given polymorphic fragment. One is a 1:1 ratio when the fragment is present in only one of the parents in heterozygous state, the other is a 3:1 ratio (presence versus absence) when the fragment originated from both parents. For the map construction, each gene was initially assigned to either a male or female linkage group depending on the donor parent of the fragments. A composite linkage map was subsequently constructed by integrating the male and female linkage groups based on common markers. The average recombination values between markers of the two parental maps were converted into map distance centiMorgans (cM) using Kosambi's mapping function (Kosambi 1944). The integrated function map of K31 is adopted to be the final function map shown in Figure 3-3 (on page 57). The genes mapped in the LH population are shown at their approximate position based on anchor RFLP markers between the two maps.

**Figure 3-3 Molecular function map of potato.** The twelve potato chromosomes are shown as a composite map based on the K31 mapping population. Genetic distances, indicated in centiMorgans on the left side of the linkage groups, are the means of the same marker intervals on maternal and paternal linkage groups (Schäfer-Pregl *et al.* 1998). Framework RFLP markers (Gebhardt *et al.* 1994) are positioned to the left side of the linkage groups and gene markers are positioned on the right side (for gene identification see Table 3-1). Relative positions of genes mapped in populations other than K31 (not connected to the vertical linkage group with a horizontal bar, see Table 3-1 for details) were inferred from reference RFLP markers. Small letters in parentheses indicate that more than one locus was identified with the same probe or PCR primer pair. Genes mapped by RFLP and STS approaches are shown in light blue and pink respectively. Map segments having QTL for TSC are shown as blue bars. QTL positions *ts(a)* to *ts(s)*, as identified by interval mapping (Schäfer-Pregl *et al.* 1998) are shown on the left side of the linkage groups.





## Results

### 3.1.4 Homology of PCR products to the reference sequences

To verify the similarity of PCR products to the genes investigated, three PCR products obtained with potato gene primers and all the six PCR products obtained with tomato gene primers and with degenerate primer pairs were sequenced. The DNA sequences were between 80 % and 100 % identical to the DNA sequences of corresponding genes in the database (Table 3-2). The variation occurred most frequently at the 3<sup>rd</sup> position of a codon followed by the 1<sup>st</sup> position while mutation at the 2<sup>nd</sup> positions was relatively less frequent. The *Pgk* gene is as an example for which a sequence was not available either in potato or in tomato. Degenerate primers were derived from consensus sequences from four different species (tobacco, Arabidopsis, spinach and wheat) and used to amplify the homologous segment in potato (see Sect. 3.1.1). A single PCR product of *Pgk* gene was obtained

(continue to page 60)

**Table 3-2 Sequence similarity between the PCR products from potato and the corresponding reference sequences**

Reference cDNA	Accession Number	Number of nucleotides compared	Similarity on nucleotide level (%)	Number of mismatches within a codon		
				1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
STP23 of potato <sup>a</sup>	D00520	660	100	0	0	0
SUS3 of potato <sup>b</sup>	U24088	710	98.5	1	0	5
SSSIII of potato <sup>c</sup>	X94400	333	100	0	0	0
GLO of tomato <sup>d</sup>	X92888	118	96.6	0	1	3
ENO of tomato <sup>e</sup>	X58108	334	97.0	2	1	7
NDPK of tomato <sup>f</sup>	X75324	390	94.4	5	2	15
PGK of tobacco <sup>g</sup>	Z48977	428	95.6	3	1	15
RCA of tomato <sup>h</sup>	AF037361	278	98.5	1	1	2
ANT of (I)	X65549	428	81.5	22	7	50
Arabidopsis <sup>i</sup> (II)		549	82.7	18	2	75

<sup>a</sup> STP23 =  $\alpha$ -Glucan phosphorylase

<sup>b</sup> SUS3 = Sucrose synthase 3

<sup>c</sup> SSSIII = soluble starch synthase III

<sup>d</sup> GLO = Glycolate oxidase

<sup>e</sup> ENO = Enolase

<sup>f</sup> NDPK = Nucleoside diphosphate kinase

<sup>g</sup> PGK = phosphoglycerate kinase

<sup>h</sup> RAC = Rubisco activase

<sup>i</sup> ANT = adenylate transporter (two PCR fragments I and II were sequenced)

## Results

1	50
PCR	TGCTATTGTG GGTGGTTCAA AGGTTTCATC CAAGATTGGA GTGAT <b>T</b> GAAT
cDNA	TGCTATTGTG GGTGGTTCAA AGGTTTCATC CAAGATTGGA GTGAT <b>C</b> GAAT
51	
PCR	CAC <b>T</b> TT <b>C</b> TGGA GAAATGTGAT ATATTGCTTT TGGGTGGAGG AATGATCTTT
cDNA	CAC <b>T</b> TT <b>T</b> TA <b>A</b> GA GAAATGTGAT ATATTGCTTT TGGGTGGAGG AATGATCTTT
101	
PCR	AC <b>T</b> TTCTACA AGGCTCAGGG TCTTT <b>C</b> AGTT GGTT <b>C</b> CTCCT TGGTTGAGGA
cDNA	AC <b>T</b> TTCTACA AGGCTCAGGG TCTTT <b>C</b> AGTT GGTT <b>C</b> CTCCT TGGTTGAGGA
151	
PCR	AGACAACTA GA <b>G</b> CT <b>T</b> GC <b>A</b> A CATCACTCCT <b>GGAA</b> AAGGCC AAGGC <b>AAA</b> AG
cDNA	AGACAACTA GA <b>A</b> CT <b>C</b> GC <b>T</b> A CATCACTCCT <b>AGAG</b> AAGGCC AAGGC <b>G</b> AAAG
201	
PCR	GAGTCAGTCT CTTGTTACCA TCTGATGTTG TGATTGCAGA TAAATTTGCT
cDNA	GAGTCAGTCT CTTGTTACCA TCTGATGTTG TGATTGCAGA TAAATTTGCT
251	
PCR	CCTGATGC <b>(C/A)</b> A ACAGCAAGGT <i>TTGCATTCTA AGCTTTTCTC ATATAAACCT</i>
cDNA	CCTGATGC <b>A</b> A ACAGCAAG-- -----
301	
PCR	<i>ATCTAACCTG AGAGCTTTCA TCTCTTGAGA TTCTTTAGAC TTGGATCTTG</i>
351	
PCR	<i>CATCTGTACT GTAATTGACA CTTAGTATCA GAATTTGTTA CTTATGGATT</i>
401	
PCR	GTGTTGAAAA TACAATTTTG TTTTG <b>G</b> TTAT TCCAGATTGT <b>GCCG</b> GCATCT
cDNA	----- --ATTGT <b>GCC<b>T</b></b> GCATCT
451	Ile
PCR	GCTATCCCAG <b>A</b> TGGTTGGAT GGGG <b>(C/T)</b> T <b>A</b> AGAC ATTGGAC <b>CT</b> G ACTCT <b>AT</b> CAA
cDNA	GCTATCCCAG <b>A</b> <b>C</b> GGTTGGAT GGGG <b>T</b> <b>TG</b> AGAC ATTGGAC <b>CA</b> G ACTCT <b>GT</b> TAA
501	Val
PCR	GACTTTCAA <b>(T/C)</b> GATGC <b>T</b> TTGG ATACCACAAA AACAGTGATC TGGAAATGGAC
cDNA	GACTTTCAA <b>C</b> GATGC <b>C</b> TTGG ATACCACAAA AACAGTGATC TGGAAATGGAC
541	Thr
PCR	CTATGGGGGT GTTTGAATTT GACAAGTTTG CT <b>ACT</b> GGAAC <b>G</b> GAGG

**Figure 3-4 Comparison of the nucleotide sequence between potato PCR product of *Pgk* gene and one of the reference cDNA sequences of the gene (tobacco).** Different nucleotides are highlighted in bold and the corresponding triplet codon is underlined. The nucleotide sequence of the intron is in italics. The alternative nucleotides at the same positions of the two alleles (SNP) are shown in parentheses. The three-letter code of amino acids was given for the non-synonymous changes.

## Results

(continued from page 58)

using genomic DNA of T710 (parent line of population LH) and the detailed sequence comparison between the PCR products and one of the reference gene sequences from tobacco is shown in Figure 3-4 (on page 59). Several features were revealed from the sequence comparison. First, the PCR product of the *Pgk* gene showed 95.6% similarity at the nucleotide level to the tobacco reference sequence. Most of the variations were point mutations and occurred at the third position of a triplet codon without changing the encoded amino acid. The only exceptions are at the positions 496-498 and 583-585. The mutations at these positions resulted in the changes of the coded amino acid from valine to isoleucine and from valine to threonine, respectively (Figure 3-4, on page 59). Second, sequence comparison of the PCR product of *Pgk* gene with the cDNA of tobacco revealed at position 268 a 167 bp intron with the characteristic feature of GT and AG motifs at each of the borders. Furthermore, sequencing from purified PCR products of the *Pgk* gene revealed directly three single nucleotide polymorphisms (SNPs) displaying a heterozygous nucleotides at positions 259, 475, and 510. Sequencing of all other PCR product revealed similar characteristics (see Section 7. Appendix)

### 3.1.5 An STS approach detected similar map positions as revealed by RFLP

Of the genes mapped, seven genes (*Sps*, *Dbe*, *Sut1*, *Pha1*, *Pha2*, *Hxk*, *GluB*) were mapped by CAPS as well as by RFLP using different mapping populations (Table 3-3). The RFLP assays were previously performed by the lab of C. Gebhardt using the cloned genes as probes with the exception of *Dbe*, for which PCR product was

**Table 3-3 CAPS markers of seven genes localised to similar map positions as mapped by RFLP**

Gene	Linkage group	Mapping approach and population	
		CAPS	RFLP
<i>Hxk</i>	III	K31	F1840
<i>Sps</i>	VII	LH	F1840
<i>Dbe</i>	XI	K31	LH
<i>Pha1</i>	III, VI	K31	F1840
<i>Pha2</i>	VII	LH	BC916 <sup>2</sup>
<i>Sut1</i>	XI	K31	BC916 <sup>2</sup>
<i>GluB</i>	I	K31	BC916 <sup>2</sup>

## Results

---

restriction enzyme *TaqI* using population F1840. These two loci were also uncovered by CAPS using population K31 in combination with *TaqI*.

### 3.1.6 Overview of the potato function map

The current potato function map consists of 66 genes (including previously mapped genes) corresponding to 82 loci distributed over all 12 chromosomes (Table 3-1, page 51-55, Figure 3-3, page 57). 56 genes are directly or potentially involved in carbohydrate metabolism and transport of which 16 are involved in starch synthesis and degradation, 9 in sucrose metabolism, 8 in transport, 8 in the Calvin cycle and photorespiration, 8 in glycolysis and the oxidative pentose phosphate pathway, and 7 in the TCA cycle. The majority of the genes were identified as a single locus, eight gene markers identified two loci (*AGPaseB*, *Inv<sub>ap</sub>*, *Pha1*, *Glo*, *Ppc*, *Fum1*, *Me*, *Ppa1*) and four gene markers detected three loci (*AGPaseS*, *rbcS*, *Ppe*, *Ldh2*).

## 3.2 Identification of single nucleotide polymorphisms (SNPs) in *Stp23*

14 potato lines with different genetic backgrounds and ploidy levels were selected to detect SNPs in *Stp23* based on sequencing of PCR products. For all these lines the PCR products of *Stp23* amplified with the primer pair shown in table 3-1 (on page 51-55) were bi-directionally sequenced. Apart from direct sequencing of the PCR products from all these lines, three lines P3, P38 and T710 were additionally cloned into the pGEM<sup>®</sup>-T vector prior to sequencing. One clone of each line was sequenced except P3 for which three clones were sequenced revealing two different alleles. The sequence comparison of all lines is shown in Figure 3-5 (on page 62-64). Sequence data obtained from PCR products and cloned fragments had readable sequences of up to 500-800 bp and 900-1100 bp, respectively, from each end. Analysis of sequences obtained from the cloned fragments revealed overlaps between the sequences of both directions, thus allowing the determination of the complete sequence of the PCR product. The PCR fragment had 2161 bases in length containing three introns with 891, 206 and 392 base pairs, respectively. For SNP identification, the low quality sequence of 360 bases in the overlapping region was excluded, only the high quality sequence region of 1801 bases was examined and compared to identify candidate SNPs. Of the nucleotide sequence examined, 653 bases are exonic while 1148 bases were intronic. (continue to page 64)

\*A H L F N H Y S S N S R F I H F T S R N

01 -----+-----+-----+-----+-----+-----+ 60

CS AGCACACTTGTTCAACCATTACAGCTCCAATTCCAGATTTCATCCATTTCACTTCTAGAAA

T710-C AGCACACTTGTTCTTACTATTACAGCTCCAATTCCAGATTTCATCCATTTCACTTCTAGAAA

Y Y

T S S K L F L T K T S H F R R P K R C F

61 -----+-----+-----+-----+-----+-----+ 120

CS CACAAGCTCCAAATTGTTCTTACCAAAACCTCCCATTTTCGGAGACCCAAACGCTGTTT

S3 CACAAGCTCCAAATTGTTCTTACCAAAACCTCCCATNTTCGGAGACCCAAACGCTGTTT

(A/T/C) (T/C)

I/T/F/S/L/P

H V N N T L S E K I H H P I T E Q

121 -----+-----+-----+-----+-----+-----+ 180

CS CCATGTCAACAATACCTTGAGTGAGAAAATTCCACCATCCCATTACTGAACAAGGTCTCAC

T710 CCATGTCAACAATACCTNGAGTGAGAAAATTCCACCATCCCATTACTGAACAAGGTCTCAC

(T/G)

W

181 -----+-----+-----+-----+-----+-----+ 240

CS TCTTATTTTCATGTATTAATTTACTATTCTGCAATTTGTTTCAATATCAAAATTTGGTAT

S3 TCTTATTTTCATGTATTAATTTACTATTCTGCAATTTGTTTCAATATCAAAATTTGGTNT

ImpL TCTTATTTTCATGTATTAATTTACTATTCTGCAATTTGTTTCAATATCAAAATTTGGTNT

T710 TCTTATTTTCATGTATTAATTTACTATTCTGCAATNGTTTCAATATCAAAATTTGGTAT

(T/G) (A/G)

241 -----+-----+-----+-----+-----+-----+ 300

CS CGAATCCATATCATTAATATACTATTTCAATATTCAAATGTGCTCTCCAAGTAGGGTA

T710-C CGAATCCATATCATTAATATACTATTTCAATATTCAAATGTGCACTCCAAGTAGGGTA

301 -----+-----+-----+-----+-----+-----+ 360

CS GATGAACCAACATTTTATACATAAACTTTTTTCAATACGTGGACTCGTGTGTTGAAACCA

S2 GATGAACCAACATTTTATACATAAACTTTTTTCAATACNTGGA

S3 GATGAACCAACATTTTATACATAAACTTTTTTCAATACNTGGA

Taig GATGAACCAACATTTTATACATAAACTTTTTTCAATACNTGGA

45C3 GATGAACCAACATTTTATACATAAACTTTTTTCAATACNTGGA

T710 GATGAACCAACNTTTTATACATAAACTTTTTTCAATACGTGGACTCGTGTGTTGAAACCA

(T/C) (A/G)

T710-C GATGAACCAACACTTTATACATAAACTTTTTTCAATACATGGACTCGTGTGTTGAAACCA

361 -----+-----+-----+-----+-----+-----+ 420

CS TTAGATCCATCTGCCTCTACATTTTCGTGGGTTAGAAAAAAAAAGTTTGTGTTGGACCTCCC

T710-C TTAGATCCATCTGCCTCTACATTTTCGTGGGTTAGAAAAAAAAAAGTTTGTGTTGGACCTCCC

P3-C1 TTAGATCCATCTGCCTCTACATTTTCGTGGGTTAGAAGAAAAAAGTTTGTGTTGGACCTCCC

421 -----+-----+-----+-----+-----+-----+ 480

CS AATCTTGTTATTGATAAAATTTTATGTTTGTGTTAAAGCACTCTTGCAAGTATTCCGGTGACT

T710-C AATCTTGTTATTGATCATATTTTATGTTTGTGTTAAAGCACTCTTCAGTATTCCGGTGACT

481 -----+-----+-----+-----+-----+-----+ 540

CS-C TCTTATTCATTTGTACTCCCAATTTAAGTGTCTTAGTTTGACAGTTTATGATATTAATG

541 -----+-----+-----+-----+-----+-----+ 600

CS-C GAGAATTTTGAATCTTGT.....GTTTTTGAAAATTTCTGG

T710-C GAGAATTTTGAATATTGTGATGTTAAATTAAACAAGTGTGTAGTTTTTGAAAATTTCTGG

601 -----+-----+-----+-----+-----+-----+ 660

CS-C TCTTAGACTTGCCAGGTAGGATATTGAAATTGGAAGACTTACTAAATAGTACTCCCTCTG



	661	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	720
CS-C	TCTCAATTTATGTGACTTACTTTTTCTTTTTAGTCAATCCAAAAAGAATGACACATTTCT		
T710-C	TC <u>CCT</u> AATTTATGTGACTTACTTTTTCTTTTTAGTCA <u>G</u> TCCAAAAAGAAT <u>A</u> ACACATTTCT		
	661	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	720
CS-C	ATATTAAATAACAATTTA ACTATAAAAATGTTTATTTTACCCTTAATGAAATAGTTTACAG		
T710-C	ATATTAAATAACAATTTA ACTATAAAAATGT <u>C</u> TATTTTACCCTTAATGAAATA <u>A</u> TTTACAG		
	721	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	780
CS-C	CCACACAAATTTCTATCATTCGTTTTGGACCACATGTTTAAAGTCTTCCTTACTTTCTT		
	781	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	840
T710-C	AAAACTCCGTGTCGAGTCAA CTACCTCACATGAAATGAGGCGGAGGGAGTATAAAAAAG		
	841	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	900
T710-C	AGACCCGACACTTCA AAAANGTATTAAAAAGATGTATGGNGGGAGTNNTN TTAATAATGG		
	901	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	960
	TGAAGCATCAGAAACCTTG ATGTATGGATCTTAGGAATTACTTACCATGATTTATGTTTA		
	961	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	1020
T710-C	TTCCCATTGAACTTGTT GTACACGATTGTATTTCCCATTAGTTCCTTCCAATTTATTACT		
	1021	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + G G E S D L S	1080
T710-C	TACATGATTTATGTTT ATTCCCATTGCTTGACATATATATAGGTGGTGAGAGCGACCTGA		
	1081	S F A P D A A S I T S S I K Y H A E F T	1140
CS-C	GTTCTTTTGCTCCTGATGCCGCATCTATTACCTCAAGTATCAAATACCATGCAGAATTCA		
T710-C	GTTCTTTTGCTCCTGATGCCGCATCTATTACCTCAAGTATCAAATACCAT <u>A</u> CAGAATTCA T		
	1141	P V F S P E R F E L P K A F F A T A Q S	1200
CS-C	CACCTGTATTCTCTCCTG AAAGGTTTGAGCTCCCTAAGGCATTCTTTGCAACAGCTCAAA		
	1201	V R D S L L I N W N A T Y D I Y E K L N	1260
CS P3-C1	GTGTTTCGTGATTTCGCT CCTTATTAATTGGAATGCTACGTATGATATTTATGAAAAGCTGA GTGTTTCGTGATTTCGCT CCTTATTAATTGG <u>H</u> CATGCTACGTATGATATTTATGAAAAGCTGA		
	1261	M K Q A Y Y L S M E F L Q	1320
CS T710-C	ACATGAAGCAAGCGTACTATCTATCCATGGAATTTCTGCAGGTATCTCATTATTCTTACT ACATGAAGCAAGCGTACTATCTATCCATGGAATTTCTGCAGGTATCTCATTATTCT <u>C</u> ACT		
	1321	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	1380
CS T710-C	TTCTCTTTTGCTCTTTTGTATGACTGTGCAGAGTGACCTTAAATTATATCTAGTAAGAAA TTCTCTTTTGCTCTTTTGTATGACTGTGCAGAGTGACCTTAAATTATATCTA <u>A</u> TAAGAAA		
	1381	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	1440
CS T710-C	TTAATCCGTTTGATATTTGCTGACAAATTAGACTGTATATTTACTGTTACACATGAGAGT TTAATCCGTTTGATATTTGCTGACAAATTA <u>T</u> ACTGTATATTTACTGTTACACATGAGAGT		
1441	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	1500	
CS	TTCTGAAATTTATGCACAA ACACTCTTTGCGAGCTGAGTTTAATTTATGCCCTATCTAC		
	1501	G R A L L N A I G N L E L T G A F A	1560
CS	TGTTTAGGGTAGAGCATTG TTAAATGCAATTGGTAATCTGGAGCTTACTGGTGCATTTCG		

	E	A	L	K	N	L	G	H	N	L	E	N	V	A	S	Q	
CS	1621	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	1680														
		GGAAGCTTTGAAAAACCTTGGCCATAATCTAGAAAATGTGGCTTCTCAGGTCAGTG TGAC															
CS	1741	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	1800														
		TTTTATTTCACGCATCAGAAAAAACCTAGTTTCCAAGTTCTTCTGTGTTTTACTCAACA															
CS	1801	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	1860														
Prio		ACAGTTGGACTATGTGTATTCATTTGTGACTTTGTGCATCGACTGTTATAA T TCAGAA															
Taig		ACAGTTGGACTATGTGTATTCATTTGTGACTTTGTGCATCGACTGTTATAA N TCAGAA															
		( <u>T/C</u> )															
T710-C		ACAGTTGGACTATGTGTATTCATTTGTGACTTTGTGCATCGACTGTTATAATAT A CAGAA															
CS	1921	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	1980														
T710-C		GCCATACATGCCATACTTGTTGTGTCTTTATGTTTTTATTTGGAGTTGACTACTTGTTAA G															
P3-C1		GCCATAC G TGCCATACTTGTTGTGTCTTTATGTTTTTATTTGGAGTTGACTACTTGTTAA															
		E P D A A L G N G G L G R															
CS	2041	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	2100														
S3		TTTTTGAATAATTATTGACAGGAACCAGATGCTGCTCTTGGAATGGGGGTTTGGGACGG															
Taig		TTTTTGAATAATTATTGACAGGAN CCAGATGCTGCTCTTGGAATGGGGGTTTGGGACGG															
		( <u>A/G</u> )															
T710-C		TTTTTGAATAATTATTGACAGGA GCAGATGCTGCTCTTGGAATGGGGGTTTGGGACGG															
		L A S C F L D S L A T L N Y P A W G Y G															
CS	2101	- - - - - + - - - - - + - - - - - + - - - - - + - - - - - +	2161														
		CTTGCTTCCTGTTTTCTGGACTCTTTGGCAACACTAAACTACCCAGCATGGGGCTATGGA															

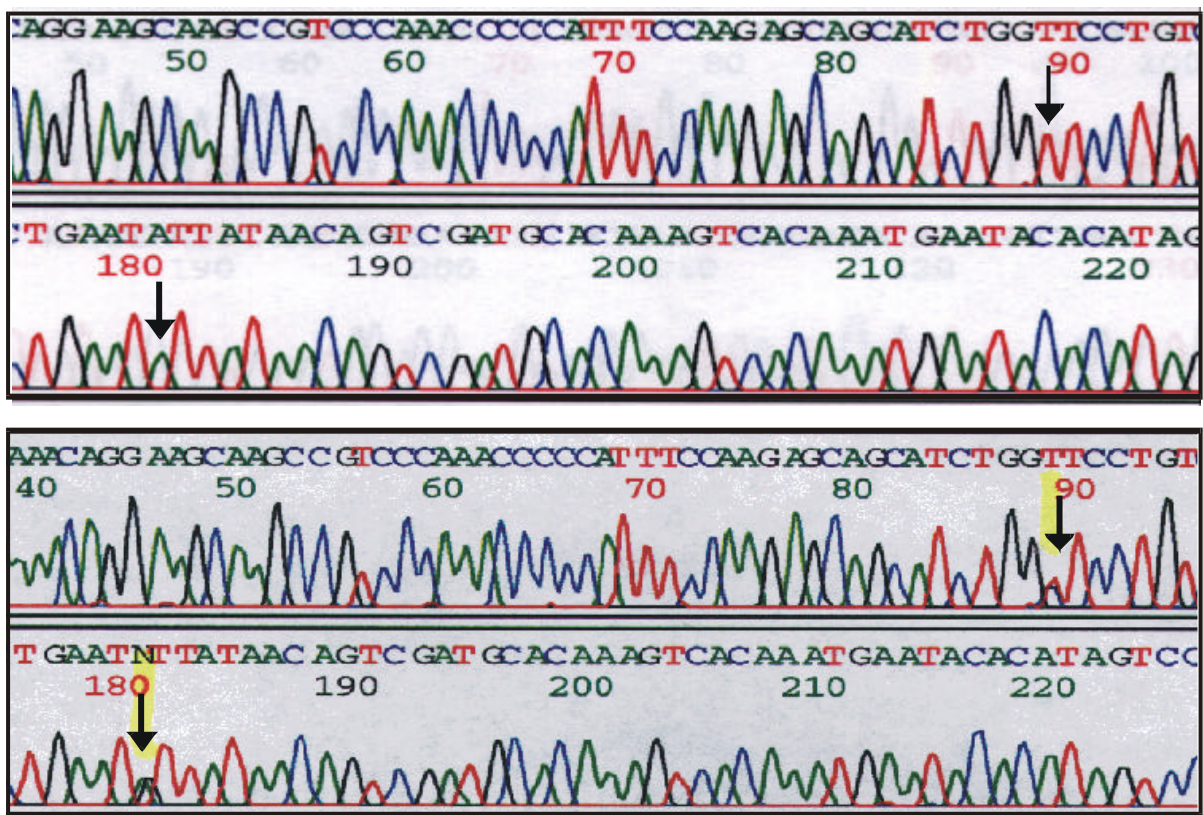
(continued from page 61)

Sequence data obtained from PCR products was first examined for different bases at single positions, then sequences were aligned and each vertical slice of the resulting assembly was examined for disagreement. Any such slice was considered a candidate SNP. Sequencing directly from PCR products allowed the identification of multiple or alternative bases at the same positions within individuals. 8 bi-allelic SNPs were detected in diploid and tetraploid potato (Figure 3-5, Figure 3-6, on page 65). In one instance, a tri-allelic SNP was found in tetraploid potato line S3 (Figure 3-6, on page 65). Sequence comparison between different lines revealed single nucleotide differences between the alleles which compose the potential SNP at that position. In three instances, the alternative base observed in different lines was

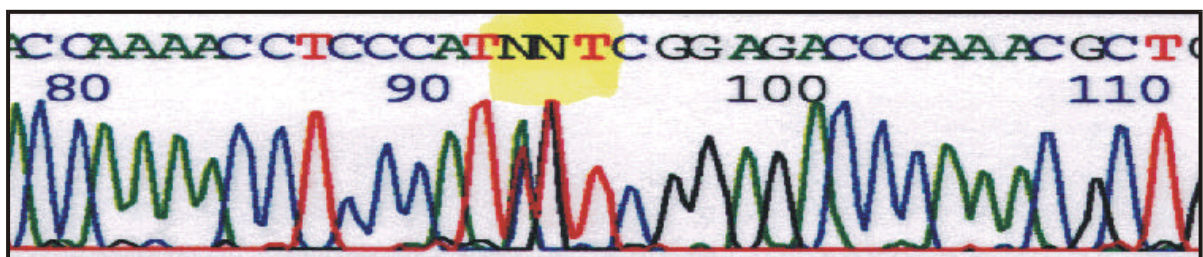
## Results

also observed within individual lines at the same position (position 314, 340 and 2065, Figure 3-5, on page 62-64). In all, 31 positions were detected which could potentially be SNPs for *Stp23* in the region examined, representing 1.7 % (31/1801) on average. 8 were in the coding region (cSNPs) whereas 23 were in introns with nucleotide diversity of 1.2% (8/653) and 2% (23/1148), respectively. Of the 8 potential cSNPs, 6 resulted in non-synonymous changes. One allele of T710, a wild species (*S. chacoense*), showed remarkably high variability as compared to the others with 20 disagreements with consensus being observed in the allele. When considering transitions versus transversions of SNPs, the frequencies of the two types

A



B



**Figure 3-6** Partial electropherogram of nucleotide sequence data from a PCR product of gene *Stp23* showing the different types of SNP. (A) bi-allelic SNP (indicated by the arrows): Homozygote (potato line Hands, upper panel) T vs. heterozygote (potato line Taiga, lower panel) C/T at position 89 and A vs. G/A at position 181; (B) tri-allelic SNP (marked yellow) in tetraploid potato line S3 at position 93.

of SNP were different. The majority of SNPs (19) involved a transition (C/T or G/A) while the remaining SNPs were transversions (C/A or G/T, C/G or G/C, T/A or A/T).

### 3.3 Molecular tagging of QTL in tetraploid potato by fAFLP markers

#### 3.3.1 Genotyping of the Z3 family by an automated fAFLP system

##### 3.3.1.1 Screening primer combinations for fAFLP

Efficient genotyping by AFLP depends on the restriction enzymes and primer combinations used. On one hand, restriction enzymes that generate fragments evenly distributed over a size of 50-500 bp facilitate data collection, thus it is highly desirable for AFLP marker generation. On the other hand, the degree of polymorphism between individuals may be variable with different primer combinations. Generation of a high degree of polymorphism between individuals could maximise the potential use of primer pairs in genotyping. Therefore, the first step in the study has been a comparison of the efficiency of different primer sets and screening of primer combinations. To do this, 12 individuals were randomly selected from population Z3 to be tested with different primer sets and primer combinations for polymorphisms. The first round of screening was carried out to determine the primer sets from four candidate pairs of restriction enzymes including *EcoRI/MseI*, *HindIII/MseI*, *EcoRI/TaqI* and *HindIII/TaqI*.

Of the four primer sets tested, *HindIII/MseI* produced less informative fragments than either *HindIII/TaqI* or *EcoRI/TaqI*, while the latter two generated larger (data not shown) and more fragments than *EcoRI/MseI* (Table 3-4, 3-5, on page 67). However, the number of polymorphic fragments did not increase. The *EcoRI/MseI* primer set, which generated more informative and more evenly distributed AFLP fragments within the size range of 50-500 bp, was selected and more primer combinations of this primer set were tested for the efficiency of generating informative fragments (Table 3-5, on page 67). A total of 29 primer combinations were further tested for primer set *EcoRI/MseI* (Table 3-5). 15 primer pairs generated more than 25 informative fragments. Based on the number of informative fragments generated and the selective nucleotide diversity of the primers, 7 of them were selected for genotyping all the individuals of breeding population Z3. The 7 primer combinations were E-aca/M-cat, E-act/M-cag, E-act/M-cta, E-aag/M-ctg, E-acg/M-cac, E-

## Results

agg/M-caa and E-acc/M-cta, which were coded as B1, B2, B3, G1, G2, G3 and Y, respectively, for simplicity.

**Table 3-4 Number of polymorphic versus total AFLP fragments generated by primer sets *EcoRI/TaqI* and *Hind III/TaqI***

	T-gta	T-gat	T-atg	T-agt	M-acg	M-aat
H-agt		3/95		3/87	19/149	10/161
H-aga			22/129	18/208	5/98	15/177
H-act	39/209		20/216	20/205	15/130	12/197
H-ata	22/121	15/103	6/152	8/139	12/96	5/90
E-act	12/166	7/156				
E-aag	16/124	18/165				

**Table 3-5 Number of polymorphic versus total AFLP fragments generated by *EcoRI/MseI* primer sets\***

	M-ctg	M-cat	M-cag	M-cac	M-cta	M-ctc	M-caa	M-ctt
E-aag	<b>31/115</b>		9/96	24/115	5/96	15/101		
E-acg			29/126	<b>27/92</b>	19/90	14/90		8/94
E-agg			28/101		8/106	25/106	<b>38/129</b>	
E-aca	6/57	<b>26/151</b>		25/103				
E-act		19/154	<b>29/109</b>	27/128	<b>27/117</b>			
E-aac			10/110					6/99
E-agc					18/135	15/104		
E-acc			29/105		<b>40/108</b>	23/117	40/149	

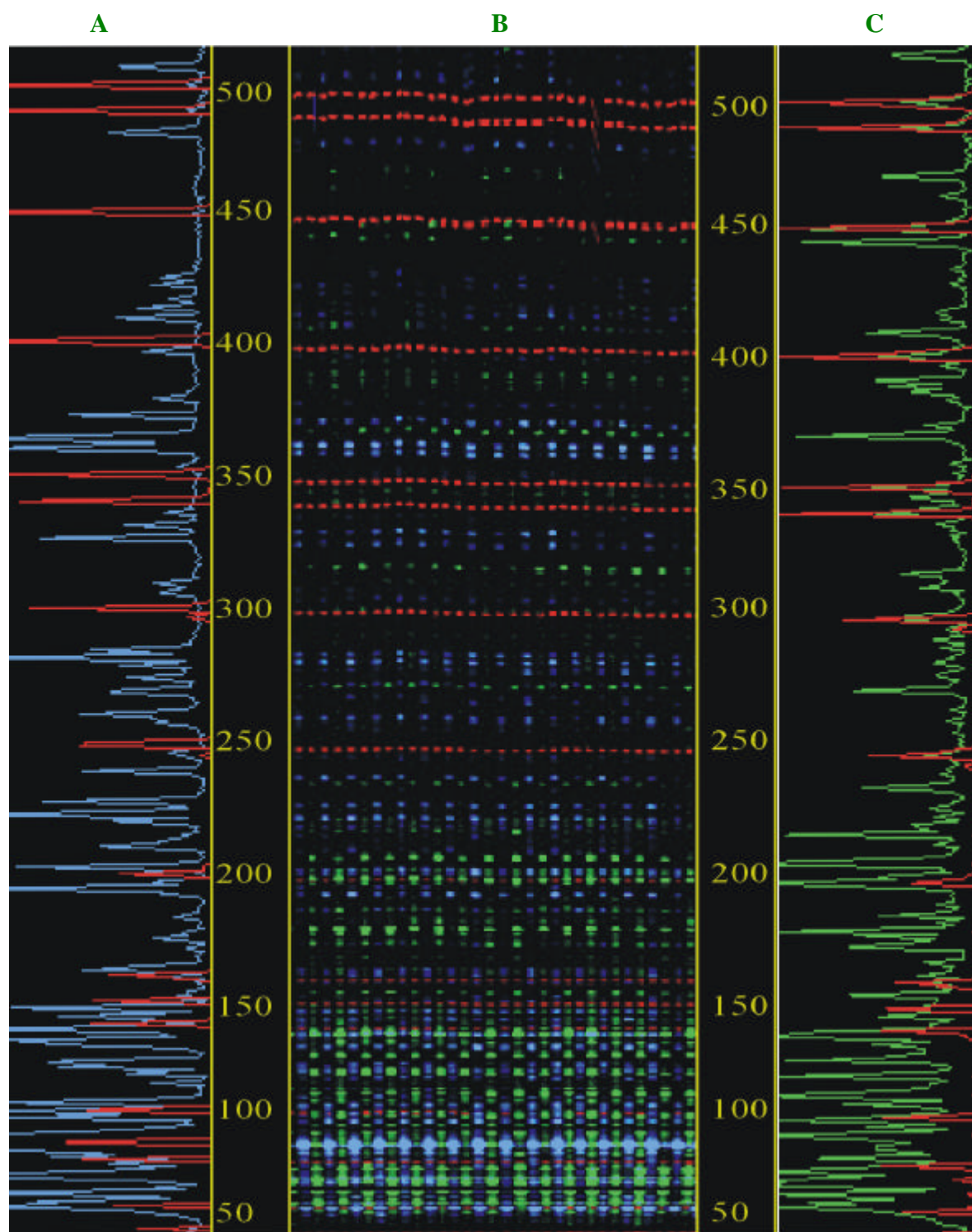
\*Primer combinations used for genotyping the Z3 population are highlighted in bold

### 3.3.1.2 fAFLP genotyping and data collection

Of the seven primer combinations selected for genotyping, selective primers E-aca and E-act were fluorescently labelled with FAM (blue), E-agg, E-acg and E-aag were labelled with JOE (green) and E-acc with TAMRA (yellow). Selectively amplified products labelled with different fluorescent dyes were alternatively loaded and resolved by electrophoresis on a 377 DNA sequencer. An example of a gel image on the 377 DNA sequencer is shown in Figure 3-7 (on page 68). Figure 3-8A (on page 69) shows fragment patterns

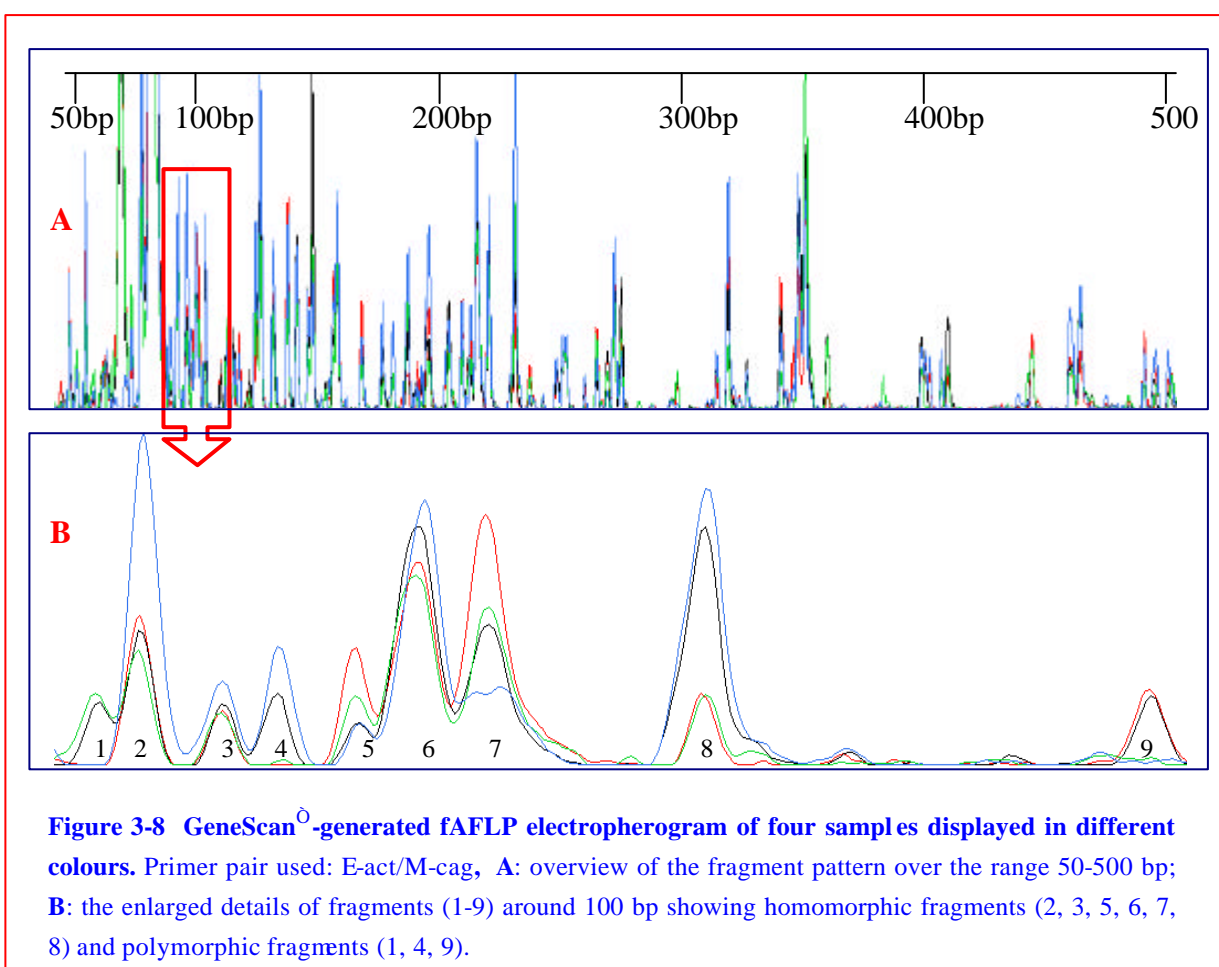
(continue to page 69)





**Figure 3-7 Electropherogram and gel view of fAFLP samples run on the ABI PRISM<sup>®</sup> 377 DNA sequencer.** Panel B shows gel image of fAFLP fragments generated with fluorescence-labelled primer pairs E-act/M-cag (blue) and E-aag/M-ctg (green). Red fragments are internal size standards with size in bp marked on the sides. Panel A and C show GeneScan<sup>®</sup>-generated electropherograms of two samples from the gel image.

## Results



(continued from page 67)

of four individuals distinguished by four colours in one panel and the Figure 3-8B shows partially enlarged details of the fragments displaying polymorphic and homomorphic fragment peaks and peak height variations. All the information in terms of both size and intensity of each fragment was collected and processed using the computer GeneScan<sup>®</sup> software. Prior to automated genotyping using the software program Genotyper<sup>™</sup>, the overall intensity of the fragments from each sample was proportionately adjusted by referring to the intensity of a pre-set reference sample. The variation of fragment intensity caused by environmental factors was therefore minimised. The final data files containing information of fragment size in base pairs and fragment intensity in peak height were created using computer software Genotyper<sup>™</sup> and exported into an *Excel* file format which was subsequently used for QTL analysis.

### *Characteristics of the fAFLP data*

The number of fragments within the size range of 50-500 bp generated with the 7 primer combinations is shown in Table 3-6 (on page 70). Between 90-150 fragments were

## Results

individually amplified with each of the 7 AFLP primer combinations. Of the fragments generated, 185 and 152 were derived from S1 and S3, respectively, while 469 descended from both parents (Table 3-6). In total, 806 different-sized fragments were scored with respect to their peak height which was detected and quantified automatically by GeneScan®. More than 200 fragments were estimated to be polymorphic based on the pre-screen data (Table 3-5, on page 67).

**Table 3-6 Number of fAFLP markers generated by seven primer combinations in population Z3 and their origins**

For a given locus in the tetraploid population, the allele dosage ranged from nulliplex to quadruplex which may contribute to the variation of the phenotypic traits, while the fragment intensity quantified by the GeneScan® software as fragment peak height reflected, to some extent, allele dosage. Thus, quantitatively (scored as peak height) and qualitatively (scored as presence and absence) scored fragments were both subjected to QTL analysis using appropriate statistical tools.

	S1	S3	Common	Total
B1	36	35	80	151
B2	24	31	53	108
B3	19	24	74	117
G1	35	19	61	115
G2	27	7	58	92
G3	27	22	80	129
Y	17	14	63	94
Total	185	152	469	806

*B1=E-aca/M-cat, B2=E-act/M-cag, B3=E-act/M-cta  
G1=E-aag/M-ctg, G2=E-acg/M-cac, G3=E-agg/M-caa  
Y=E-acc/M-cta*

### 3.3.2 Analysis of phenotypic traits and phenotypic data\*

150-200 individuals of population Z3 were evaluated for five phenotypic traits including tuber starch content (TSC), reducing sugar content (RSC), foliage earliness (FE) development (DEV) and tuber yield (TY). The phenotypic data for 1997, which was created from scoring a single plant individual (no replication), was not subjected to analysis separately but was merged with the data of 1998 for which the scores resulted from 5 different plants of each clone (Strahwald, 1999, personal communication). The combined data were used for further analysis. Heritability and correlation coefficients of the traits were calculated and are shown in Table 3-7 based on 1998 data.

(continue to page 72)

\*Phenotypic data were evaluated in field trials at two locations (Gransebieth and Windeby) over three years (1997, 1998, and 1999) by the potato breeding company SaKa-Ragis (Dr. Josef Strahwald). They are summarized here because the phenotypic data was an essential part of the QTL analysis.



## Results

**Table 3-7 Heritability and correlation coefficient of the phenotypic traits evaluated\***

	$H^2$	$r$		
		DEV	FE	TSC
FE	0.65	0.27		
TSC	0.59	0.17	0.49	
RSC	0.62		-0.26	-0.62
TY	0.14		-0.18	
DEV	0.32			

FE= foliage earliness  
TSC= tuber starch content  
RSC= reducing sugar content  
TY= tuber yield  
DEV= development  
 $H^2$ = broad-sense heritability  
 $r$ = correlation coefficient  
significance level: 0.01  
Correlation coefficients, which were not significant at this level, are not shown in the table.

**Figure 39 Frequency distribution of phenotypic scores for**

**A: reducing sugar content**

(% fresh weight)

**B: foliage earliness (1= very early, 9= very late);**

**C: tuber starch content**

(% fresh weight)

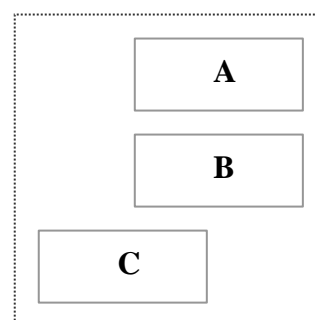
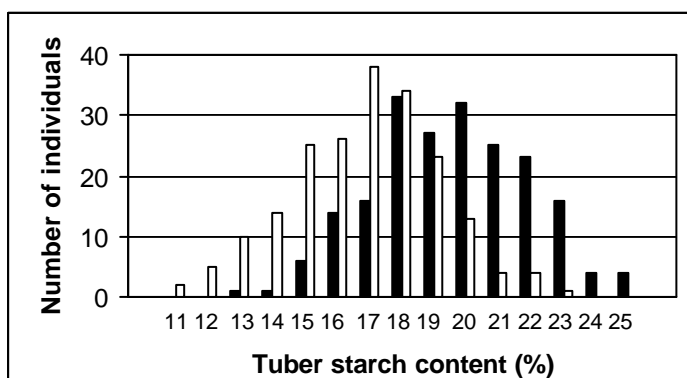
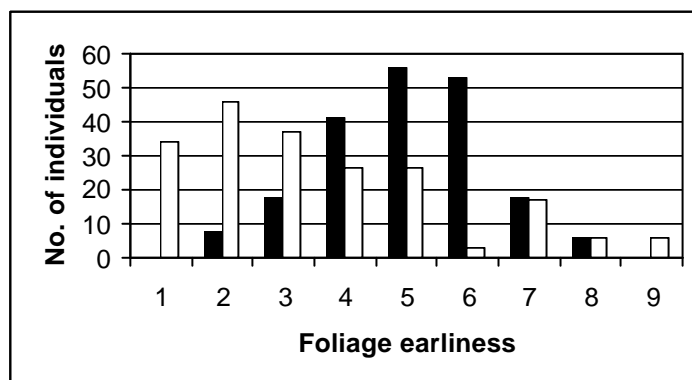
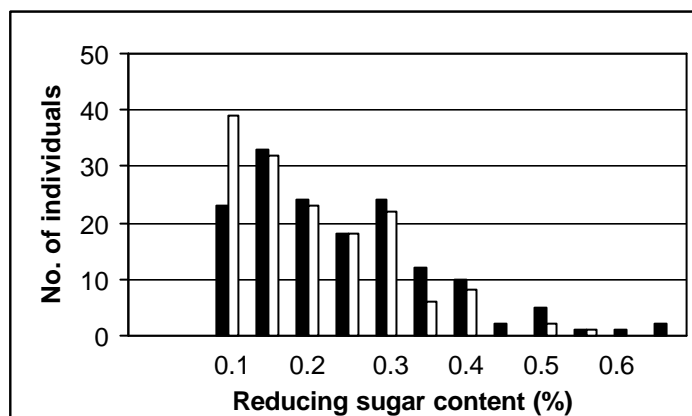
Based on the data collected from trial 1998 at two locations,

Gransebieth (blank column) and

Windeby (black column) using

population Z3.

200 individuals were evaluated for foliage earliness and tuber starch content, whereas 150 individuals were assessed for reducing sugar content.



## Results

---

*(continued from page 70)*

Of the five traits investigated, significant correlations were revealed among the traits (Table 3-7, on page 71). It should be noted that TSC, RSC and FE showed significant correlations and relatively high correlation coefficients between each other. TSC had a significant negative correlation with RSC ( $r=-0.62$ ,  $P<0.01$ ) and positive correlation with FE ( $r=0.49$ ,  $P<0.01$ , the higher the TSC, the later the foliage matures). There are also significant correlations between RSC and FE ( $r=-0.26$ ,  $P<0.01$ , the higher the reducing sugar content, the earlier the foliage matures).

Heritability of different traits was quite variable in the population studied. TSC, RSC and FE showed high heritability, while tuber yield and development had rather lower heritability (Table 3-7, on page 71). These observations indicated that the phenotypic variation observed for the former three traits resulted mainly from genetic variation, whereas, the latter two were largely caused by environmental factors. Therefore, further studies were only concentrated on the three high heritability traits TSC, RSC and FE.

Based on the method used for evaluating the three traits (see Sect 2.14), it is obvious that the phenotypic data for FE and RSC were categorical since they were based on or converted from a 1-9 reference standard scale. Thus both data sets are discontinuous. The data for TSC is quantitative, continuous and normally distributed (Figure 3-9C, on page 71) which was indicated by a Kolmogorov-Smirnov test ( $p > 0.2$ ). TSC ranged from 13.6 % to 25.8 % in Windeby and 7.1 % to 22.4 % in Gransebieth with mean values of 20.21% and 16.34 %, and standard deviation of 2.36 and 2.29, respectively. The phenotypic distribution of the three traits are graphically exhibited in histograms of Figure 3-9 (on page 71). The distributions were consistent with the assumed polygenic inheritance of the traits considered. According to the nature of the phenotypic data, parametric and non-parametric methods of statistical analysis were applied to detect marker-trait associations for TSC, RSC and FE.

### 3.3.3 Identification of QTL associated with fAFLP fragments

The QTL analysis was performed in collaboration with Dr. Schäfer-Pregl (MPIZ). The data collected from the two different locations were analysed to determine the significance and the amount of phenotypic variance explained by fAFLP markers. Statistical analysis as described in Sect. 2.15 was applied to identify the associations of phenotypic traits with fAFLP

*(continue to page 74)*

## Results

**Table 3-8 fAFLP markers showing associations with QTL for TSC as detected by the t-test and regression analysis**

fAFLP marker	LG <sup>a</sup>	No -/+ <sup>b</sup>	$\chi^2$ -test <sup>c</sup>	Ori. <sup>d</sup>	t-test				Regression analysis			
					Gransebieth		Windeby		Gransebieth		Windeby	
					<i>P</i> <sup>e</sup>	R <sup>2f</sup> (%)	<i>P</i> <sup>e</sup>	R <sup>2f</sup> (%)	<i>P</i>	R <sup>2f</sup> (%)	<i>P</i>	R <sup>2f</sup> (%)
B3-409		138/51		SI	0.00017+	6.0	0.0018+	5.4	0.0075	3.9	0.0023	5.0
B1-213		158/27		S3	0.014-	4.4	0.0053-	5.6	0.011	3.4	0.032	2.5
B1-214		36/109	1:3	C	0.015-	3.6	0.0051-	4.7	0.0014	5.3	0.0047	4.2
B2-429		138/42		S3	0.025-	3.1	0.04-	2.8	0.02	3.0	0.017	3.2
G1-169		136/32		S3	0.0076-	3.2	0.014-	3.0	0.018	3.3	0.035	2.6
G1-435		116/51		SI	0.022+	3.1	0.0046+	4.5			0.018	3.3
B2-306		15/146	1:11	C	0.041-	3.5	0.026-	4.2			0.022	3.2
B1-174	V	21/174	1:11	C	0.045+	2.7	0.033+	1.5				
B3-134		143/40		SI	0.022+	1.6	0.017+	2.8				
B3-397		137/45		SI	0.0023+	5.1	0.041+	2.2				
B1-124		36/158		C	0.033-	2.1	0.026-	2.7				
G3-393	I	133/46		SI	0.010-	4.7	0.024-	3.2				
B1-244		137/43		S3	0.018-	3.6			0.009	3.8		
B2-428		57/122		C	0.011+	4.2			0.014	3.4		
B3-260	II	135/52		SI	0.011+	3.1			0.01	3.6		
G1-325		118/24		C	0.016+	3.3			0.024	2.7		
G2-187		49/137	1:3	C	0.0037-	4.2			0.0066	4.1		
G3-68		90/95		C	0.015-	3.2			0.024	2.7		
G3-101		75/110		C	0.0053-	4.1			0.027	2.6		
B1-316	III	67/119		C			0.0029-	4.7			0.0078	3.8
B2-168	V	104/77		SI			0.048+	2.2			0.043	2.3
B2-352		116/65		SI			0.01-	3.8			0.033	2.5
B3-121	IX	62/130		C			0.0035+	3.8	0.034	2.4	0.0003	6.9
G1-200		140/49		SI			0.032-	2.3			0.0094	3.5
G3-163		49/132	1:3	C			0.028-	2.5			0.032	2.5
G3-342	III	61/116		S3			0.043+	2.6			0.01	3.7
G1-156		20/171	1:11	C			0.045-	2.1	0.0066	3.8		
Y-167		57/97		SI	0.022-	3.3	0.043-	2.5				
B1-166	I	100/94	1:1	S3	0.0069-	3.7						
B2-76	V	90/95		C			0.0011-	5.7				
G3-490		49/122	1:3	C					0.002	5.2	0.022	2.9
B1-107		191		C					0.013	3.2	0.018	2.9
G1-184		185		C					0.013	3.3	0.0022	5.0
G1-109		190		C					0.0006	6.0		
G1-324		82/93	1:1	SI					0.0049	4.2		
G3-437		48/111		C					0.006	4.5		
B2-173	XII	186		C							0.0021	5.0
B3-120		183		C							0.0057	4.1
G1-330		49/140	1:3	C							0.0017	5.1
G3-429		107/76		C							0.0058	4.1

<sup>a</sup> LG= linkage group, inferred from map positions of common fAFLP markers in the reference diploid mapping population K31.

<sup>b</sup> No. -/+, number of individuals analyzed showing absence/presence at the marker locus.

<sup>c</sup> 1:N, segregation type confirmed by goodness-of-fit test.

<sup>d</sup> Ori., origin of the marker with 'C' representing common marker from both parents.

<sup>e</sup> *P* values followed by plus or minus indicate the direction of the tagged QTL effects with '+' being increasing and '-' being decreasing TSC.

<sup>f</sup> R<sup>2</sup>= R square value in percent indicating the proportion of variance explained by the marker genotype.

## Results

---

*(continued from page 72)*

markers. Briefly, parametric tests, a *t*-test and regression analysis, were applied to data of TSC, while the non-parametric Mann-Whitney U-test was used for association analysis between markers and traits for data of RSC and FE. The existence of a fAFLP-tagged QTL was accepted when the marker was consistently detected to be associated with one trait in both locations, or with two or more different traits at a threshold of  $P < 0.05$ . Fragments for which association was detected in only one location with only one trait was considered of potential interest only when  $P < 0.01$ . The fAFLP fragments, which meet the above criteria, are presented in the Sect. 3.3.3.1.

### **3.3.3.1 Association of fAFLP fragments with tuber starch content (TSC)**

Association analysis using the *t*-test and linear regression resulted in a total of 40 fAFLP markers putatively associated with QTL for TSC with the threshold chosen. A nomenclature for fAFLP was adopted by starting with a primer combination code followed by the fragment size detected. The fAFLP markers, magnitude and direction of QTL effects are summarized in the order of their reproducibility in Table 3-8 (on page 73). Among the 40 fAFLP markers as identified by *t*-test and regression analysis, 5 markers (B3-409, B1-213, B1-214, B2-429, G1-169) were consistently detected at two locations and with two different tests; 3 markers (G3-490, B1-107, and G1-148) were repeatedly detected but only with regression analysis; 7 markers (G1-435, B2-306, B1-174, B3-134, B3-397, B1-124, and G3-393) were reproducible at two locations using the *t*-test but not detected or detected only at one location using regression analysis. 15 markers were reproducibly detected using both the *t*-test and regression analysis in one of the locations. Others were detected only in a single test. Of the 12 markers consistently detected at both locations using the *t*-test, markers B3-409, G1-435, B1-174, B3-134 and B3-397, which all originated from parent S1, showed a positive effect on TSC whereas other markers descending from both parents or from S3 only, showed negative effects on TSC. This observation suggested that the S1 allele at these marker loci was associated with an increase in TSC. These markers, by ANOVA, explained 1.5-6 % of the phenotypic variation individually. The most significant effect on TSC was associated with the marker B3-409. The allelic variation at this marker locus accounted for up to 6 % of the phenotypic variation. Given the broad-sense heritability of 59 % for TSC, this marker explained 10 % of the total genetic variance.

## Results

### 3.3.3.2 Association of fAFLP fragments with reducing sugar content (RSC) and foliage earliness (FE)

Due to the discontinuous nature of the phenotypic data for the both traits, the non-parametric Mann-Whitney U-test was applied to the analysis for the marker-trait association. The results were shown in Table 3-9.

#### *fAFLP markers associated with RSC*

Over two locations, a total of 25 markers were associated with RSC (Table 3-9). 11 of these markers were consistent across two locations at the 5% or 1% significance level, the others were detected at only one location at the 1% level. The majority of the markers (10) tagged alleles of the parent S3 which contributed to increased RSC. (continue to page 76)

**Table 3-9 fAFLP markers showing associations with QTL for RSC and FE as detected by the Mann-Whitney U-test**

fAFLP marker	LG <sup>a</sup>	No -/+ <sup>b</sup>	$\chi^2$ -test <sup>c</sup>	Ori. <sup>d</sup>	Reducing sugar content				Foliage earliness			
					Gransebieth		Windeby		Gransebieth		Windeby	
					P <sup>e</sup>	R <sup>2f</sup> (%)	P <sup>e</sup>	R <sup>2f</sup> (%)	P <sup>e</sup>	R <sup>2f</sup> (%)	P <sup>e</sup>	R <sup>2f</sup> (%)
B3-117		18/117		C	0.009-	3.9	0.0009-	12.6				
B3-413		40/96	1:3	C	0.019-	2.4	0.002-	5.7				
B3-69		75/70	1:1	S1	0.013-	3.7	0.027-	2.8				
B2-123		27/113	1:3	C	0.044+	1.8	0.017+	2.0				
B3-360		42/103	1:3	C	0.026+	3.4	0.021+	3.6				
G2-386		73/44		S3	0.003+	7.2	0.002+	5.2				
G3-102		9/128	1:11	C	0.034+	2.7	0.016+	3.3				
Y-452		40/91	1:3	C	0.037+	3.8	0.045+	3.1				
B1-330		94/47		S3	0.009+	3.6	0.006+	6.1	0.041-	1.6		
B3-215	V	84/60		S3	0.0004+	6.6	0.0001+	11.3	0.043+	1.9		
G3-79		82/58		S3	0.005+	7.8	0.029+	2.9	0.005+	4.6	0.001+	5.2
B1-158	II	42/149	1:3	C					0.043+	2.8	0.027+	2.3
B1-225		147/33		S3					0.038+	2.4	0.014+	4.1
B2-67		105/81	1:1	S1					0.039+	3.2	0.042+	2.1
B2-153		94/88	1:1	S3					0.046+	1.6	0.009+	4.0
B2-185		84/99	1:1	S3					0.0008+	5.3	0.019+	2.7
B2-314		128/52		S1					0.008+	3.2	0.003+	4.6
B3-53		5/188	1:35	C					0.019+	2.1	0.019+	3.0
B3-100	XII	10/182	1:11	C					0.011+	2.8	0.001+	6.4
B3-159		34/151	1:3	C					0.035+	1.5	0.026+	3.3
B3-305		85/99	1:1	S3					0.036+	2.3	0.019+	3.3
B3-327		100/90	1:1	S3					0.021+	3.6	0.035+	2.3
B3-352		138/51		S3					0.027+	2.7	0.051+	2.2
B3-425		132/44		S1					0.006+	3.9	0.064+	2.3
G3-129		29/156		C					0.002+	5.0	0.004+	4.5
G3-202	XII	115/65		S3					0.014+	2.2	0.026+	2.5
Y-393		43/128		S1					0.024+	2.0	0.030+	3.0
G3-54	IX	57/81		S1			0.034-	4.8	0.027-	2.8	0.010-	4.7
B1-241		22/169	1:5	S3					0.003-	5.7	0.003-	5.0
B2-58		58/124		C					0.004-	4.2	0.0005-	6.3
B2-76	V	90/95		C					0.0004-	5.9	0.0005-	6.9

## Results

**Table 3-9 (continued)**

fAFLP marker	LG <sup>a</sup>	No -/+ <sup>b</sup>	$\chi^2$ - test <sup>c</sup>	Ori. <sup>d</sup>	Reducing sugar content				Foliage earliness			
					Gransebieth		Windeby		Gransebieth		Windeby	
					P <sup>e</sup>	R <sup>2f</sup> (%)	P <sup>e</sup>	R <sup>2f</sup> (%)	P <sup>e</sup>	R <sup>2f</sup> (%)	P <sup>e</sup>	R <sup>2f</sup> (%)
G3-206		44/138		S3					0.017-	2.7	0.017-	3.6
Y-459		106/65		S3					0.018-	2.0	0.036-	2.8
B1-182		26/115	1:3	C			0.044+	2.8	0.034+	2.7		
B1-330		68/73	1:1	S3	0.002+	5.1			0.036-	1.9		
B3-330		72/70	1:1	S3	0.023+	3.6			0.049-	1.6		
B1-77	XI	37/111	1:3	C	0.005+	5.4						
B1-102		51/95		S3	0.002+	7.3						
B1-455		90/52		S3	0.004+	6.7						
B3-141	V	70/76	1:1	S1	0.005-	5.3						
B3-211	I	63/81		C	0.007+	6.3						
G2-409		64/66	1:1	S1	0.006+	4.1						
B1-140		8/140	1:11	C			0.009+	4.0				
B1-261	IV	78/64	1:1	S1			0.005+	5.4				
B1-423		66/77	1:1	S3			0.004+	4.9				
B2-397		89/47		S3			0.004-	5.0				
G1-339		40/104	1:3	C			0.007+	5.1				
Y-111		13/120	1:11	C			0.009+	4.7				
B1-174	V	21/174	1:11	C					0.006+	3.5		
B1-322		48/138	1:3	C					0.007+	4.2		
B2-158		104/77		S1					0.001+	4.5		
G2-67		115/78		C					0.007-	2.7		
B2-68		21/164	1:5	S3							0.006+	4.0
G1-132		137/51		S3							0.009+	4.1
G2-80		76/110		C							0.002-	4.1
G2-120		35/157		C							0.004+	4.7
Y-83		83/92	1:1	S3							0.003-	4.5

<sup>a</sup> LG= linkage group, inferred from map positions of common fAFLP markers in the reference diploid mapping population K31.

<sup>b</sup> No. -/+, number of individuals analyzed showing absence/presence at the marker locus.

<sup>c</sup> 1:N, segregation type confirmed by goodness-of-fit test.

<sup>d</sup> Ori., origin of the marker with C representing common marker from both parents.

<sup>e</sup> P values followed by plus-minus the indicating the direction of the tagged QTL effect with '+' being increasing and '-' being decreasing RSC or FE.

<sup>f</sup> R<sup>2</sup>= R square value in percent indicating the proportion of variance explained by the marker genotype.

*(continued from page 75)*

The only exception was marker B2-397, which decreased RSC. Only four markers were derived from the S1 parent, with two increasing RSC and two reducing RSC. The most significant markers are B3-117, B3-413, B1-330, B3-215, and G3-79. Individually, these markers each explained between 2.4 to 12.6% of the phenotypic variation among individuals for RSC.

### ***fAFLP markers associated with FE***

Compared to RSC, more fAFLP markers (32) were detected to be significantly associated with FE (Table 3-9). Effects of 23 markers were reproducibly detected (*continue to page 79*)

## Results

**Table 3-10 fAFLP markers showing associations with multiple phenotypic traits**

fAFLP marker	LG <sup>a</sup>	No -/+ <sup>b</sup>	$\chi^2$ -test <sup>c</sup>	Ori. <sup>d</sup>	Tuber starch content				Reducing sugar content				Foliage earliness			
					Gransebieth		Windeby		Gransebieth		Windeby		Gransebieth		Windeby	
					<i>P</i> <sup>e</sup>	<i>R</i> <sup>2f</sup> (%)	<i>P</i> <sup>e</sup>	<i>R</i> <sup>2f</sup> (%)	<i>P</i> <sup>e</sup>	<i>R</i> <sup>2f</sup> (%)	<i>P</i> <sup>e</sup>	<i>R</i> <sup>2f</sup> (%)	<i>P</i> <sup>e</sup>	<i>R</i> <sup>2f</sup> (%)	<i>P</i> <sup>e</sup>	<i>R</i> <sup>2f</sup> (%)
B3-409		138/51		SI	0.00017+	6.0	0.0018+	5.4			0.018-	3.9				
B3-134		143/40		SI	0.022+	1.6	0.017+	2.8			0.012-	5.0				
B1-214		36/109	1:3	C	0.015-	3.6	0.0051-	4.7	0.031+	3.8						
B1-166	I	100/94	1:1	S3	0.0069-	3.7			0.050+	2.2						
B1-77	XI	49/146	1:3	C	0.027-	2.9			0.005+	5.4						
B1-390		83/59		C	0.045-	2.2			0.023+	2.5						
B2-292		20/118	1:5	SI	0.013-	3.1			0.048+	2.8						
B2-62		79/103		SI	0.021+	2.9					0.039-	3.3				
G1-339		50/139	1:3	C	0.011+	3.6					0.007+	5.1				
G1-157		116/75		SI			0.025-	2.7			0.029+	3.8			0.019-	2.6
Y-452		51/120		SI			0.027-	2.8	0.037+	3.8	0.045+	3.1				
B1-112		77/117		C			0.035-	2.3	0.022+	2.5						
B1-149	III	9/183		C			0.028-	3.1	0.031+	2.8						
B1-364		112/70		C			0.040-	2.2	0.039+	3.2						
B3-121	IX	62/130		C			0.0035+	3.8	0.046-	2.3						
B1-174	V	21/174	1:11	C	0.045+	2.7	0.033+	1.5					0.006+	3.5		
G2-187		49/137	1:3	C	0.0037-	4.2									0.010-	2.9
B2-322		140/40		SI	0.019+	3.2							0.024+	4.7		
G3-199	IV	26/151	1:5	SI	0.019-	3.0									0.032-	1.8
B1-244		137/43		S3	0.018-	3.6									0.048+	1.9
Y-393		43/128		SI	0.021-	2.9							0.024+	2.0	0.030+	3.0
G1-200		140/49		SI			0.032-	2.3					0.016-	2.4		
B2-58		58/124		C			0.036-	2.8					0.004-	4.2	0.0005-	6.3
B3-100	XII	10/182		C			0.013+	2.2					0.011+	2.8	0.001+	6.4
B2-76	V	90/95		C			0.0011-	5.7					0.0004-	5.9	0.0005-	6.9
G1-205		103/86	1:1	SI			0.048-	2.0							0.047-	2.5
B1-242		107/76		SI			0.029-	2.6							0.015-	2.5

## Results

**Table 3-10 (continued)**

fAFLP marker	LG <sup>a</sup>	No -/+ <sup>b</sup>	Ori. <sup>c</sup>	Tuber starch content				Reducing sugar content				Foliage earliness			
				Gransebieth		Windeby		Gransebieth		Windeby		Gransebieth		Windeby	
				P <sup>d</sup>	R <sup>2e</sup>	P <sup>d</sup>	R <sup>2e</sup>	R <sup>2e</sup>	P <sup>d</sup>	R <sup>2e</sup>	P <sup>d</sup>	R <sup>2e</sup>	P <sup>d</sup>	R <sup>2e</sup>	
				(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)		
B1-330		94/47	S3					0.009+	3.6	0.006+	6.1	0.041-	1.6		
B3-215	V	84/60	S3					0.0004+	6.6	0.0001+	11.3	0.043+	1.9		
G3-79		82/58	S3					0.005+	7.8	0.029+	2.9	0.005+	4.6	0.001+	5.2
G3-54	IX	57/81	S1							0.034-	4.8	0.027-	2.8	0.010-	4.7
B1-182		26/115	1:3	C						0.044+	2.8	0.034+	2.7		
B3-330		72/70	1:1	S3				0.023+	3.6			0.049-	1.6		

<sup>a</sup> LG= linkage group, inferred from map positions of common fAFLP markers in the reference diploid mapping population K31.

<sup>b</sup> No. -/+, number of individuals analysed showing absence/presence at the marker locus.

<sup>c</sup> 1:N, segregation type confirmed by goodness-of-fit test.

<sup>d</sup> Ori., origin of the marker with C representing common marker from both parents.

<sup>e</sup> P values followed by plus-minus the indicating the direction of the tagged QTL effect with '+' being increasing and '-' being decreasing TSC or RSC or FE.

<sup>f</sup> R<sup>2</sup>= R square value in percent indicating the proportion of variance explained by the marker genotype.



## Results

---

*(continued from page 76)*

over two locations, 17 of which resulted in delayed foliage maturity (positive effect noted with '+') while 6 lead to earlier maturity. The most significant ones were fragments B2-314, G3-129, B1-241, B2-58, B2-76, with effects accounting for phenotypic variation of 3.2 % - 6.9 % individually (Table 3-9, on page 75-76). Among these 23 markers, two effects with opposite direction were observed for markers of different parental origin. S1 alleles at the marker locus of G3-54 and S3 alleles at the marker loci of B1-241, G3-206, Y-459 were responsible for earlier foliage maturity, both S1 and S3 alleles at other marker loci resulted in delayed foliage maturity.

### **3.3.3.3 Association of fAFLP fragments with multiple traits**

Of all the fAFLP markers analysed, 33 have been identified to be associated with more than one trait as detected by the t-test, regression analysis and the U-test. These markers were extracted and are shown in Table 3-10 (on page 77-78). 15 markers that were associated with TSC, also showed associations with RSC in at least one location. The most significant and reproducible ones were markers B3-409 and B3-134, which originated from S1 parent. These two markers exhibited effects of increasing TSC and decreasing RSC and thus, could be of potential use in potato breeding programs for MAS. It is worth noting that all markers had an opposite effect on the two traits, with the only exception being G1-339. In other words, markers with positive effects of increasing TSC would lead to decreased RSC or vice-versa. The results were in fairly good agreement with the phenotypic correlation relationship between the two traits as outlined in Section 3.3.2.

12 fragments showed multiple associations with TSC and FE. The most significant ones are G2-187, B2-58 and B2-76. Most frequently, the fragments with increasing effects on TSC usually lead to delayed maturity and vice-versa. The only exception was marker Y-393 for which the presence of the fragment decreased TSC and delayed maturity.

7 markers showed multiple associations with RSC and FE. The qualitative orientations of effects were variable among these markers. 3 markers (G1-157, B1-330, B3-330) with increasing effects on RSC lead to earlier maturity while 3 (B3-215, G3-79, B1-182) lead to delayed maturity. In one case, at locus G3-54, the decreased RSC was associated with earlier maturity.

One marker, G1-157 of S1 was associated with all three traits at one location (Windeby).

## Results

---

Presence of the marker decreased TSC, increased RSC and lead to earlier maturity.

### **3.3.3.4 Association analysis on 50 selected genotypes in trial 1999**

In the trial 1999, 50 genotypes selected from the Z3 population were evaluated for various traits (see Sect.2.14). In addition to TSC and FE (RSC was not evaluated), other traits namely starch yield (SY) and tuber yield (TY) were also tested. The phenotypic data of these traits were, therefore, subjected to association analysis using the non-parametric U-test as described previously. The significantly associated fAFLP markers are shown in Table 3-11 (on page 81). A total of 42 fAFLP markers showed associations with starch yield, of which 18 and 21 were detected at either Gransebieth or at Windeby, and 3 were consistently detected at both locations. Of all these, 10 and 28 markers were also detected, at the same locations, to be associated with TSC and TY, respectively. This suggested that the effects of these markers on TSC and TY contributed to their effects on SY. 10 markers, which showed significant effects on either TSC or TY at the 1 % level, did not detect as having significant effects on SY.

Of the markers showing association with the traits in trial 1999, markers highlighted in bold (G1-75, B3-397, B2-428, B3-305, B3-51, G3-111G1-169, B1-215, B3-215, G3-54, and B3-352) were also detected, at least at the 5% level, as having association with the same traits in trial 1998. It is worth noting that for markers showing association with multiple traits, the directions of effects were uniformly in accordance to the phenotypic correlations (Table 3-11, page 81). In other words, the increase in either TSC or TY resulted in the increase of SY while leading to delayed maturity and vice versa.

### **3.3.4 Analysis of the segregation ratios in the tetraploid population Z3**

Several segregation ratios were expected depending on the presence or absence of a fAFLP marker fragment in the parents. When a marker is present in only one parent, segregation ratios of 1:1 or 1:5 (absence versus presence) are expected for simplex and duplex markers, respectively. When a marker is present in both parents, segregation of 1:3, 1:11 and 1:35 (absence versus presence) are expected for simplex-simplex, simplex-duplex and duplex-duplex, respectively, without considering double reduction. Chi-square tests were performed to test for the goodness of fit to the expected segregation ratios. Markers fitting one of the expected segregation ratios were noted in Table 3-8, 3-9, and 3-10 (on page 73, 75-76, 77-78).

## Results

**Table 3-11 fAFLP markers associated with phenotypic traits as detected by the U-test based on 50 selected genotypes in trial 1999**

	Pre/abs <sup>b</sup>	Ori./LG	SY <sup>c</sup>		TSC <sup>c</sup>		TY <sup>c</sup>		FE <sup>c</sup>	
			Gra	Win	Gra	Win	Gra	Win	Gra	Win
B3-62	25/23	S3	0.012+	0.020+	0.0013+	0.008+				
G1-133	25/23	S3	0.024+	0.027+	0.033+	0.049+				
B1-144	36/12	C	0.028+	0.0005+				0.0002+		
<b>G1-75<sup>a</sup></b>	30/19	S1	0.0013+			0.036+	0.0039+			
B3-269	16/25	S3/VIII	0.0028+				0.013+			
Y-60	30/15	S1	0.013+		0.021+					
G2-237	15/33	S1	0.038+		0.0081+					
G3-231	32/15	C	0.0078-		0.0073-					
G2-238	13/31	S3	0.019-		0.0048-					
G2-108	33/17	S3	0.032-		0.028-					
G2-301	21/29	S1	0.027+			0.013+				
<b>B3-397</b>	13/34	S1	0.029+			0.016+				
G1-106	34/15	S1/XII	0.040+			0.039+				
B1-464	19/28	C	0.039+					0.0067+		
G3-236	21/26	S1	0.010+				0.034+			
<b>B3-305</b>	25/20	S3/IV	0.011+				0.032+			0.046+
B3-223	40/7	C	0.014+				0.040+			
<b>B2-428</b>	30/18	C	0.029+				0.023+			
G1-105	27/18	S1	0.031+				0.034+			
B2-274	43/5	C	0.035+				0.0076+			
B2-324	15/33	S3	0.012-				0.022-			
<b>B3-51</b>	32/16	C		0.0025+	0.0016+	0.039+		0.013+	0.016+	0.013+
G2-120	40/9	C		0.0057+				0.019+	0.029+	0.037+
B2-185	22/26	S3		0.027+	0.046+			0.048+	0.0057+	
<b>G3-111</b>	33/14	C		0.0042+				0.026+		
B1-267	41/8	C		0.019+				0.0020+		
B3-104	43/5	C		0.023+				0.0059+		0.032+
B1-78	34/15	C		0.010+				0.009+		
B2-216	31/15	C		0.010+				0.013+		
G1-115	19/23	S1		0.013+				0.019+		
G2-164	18/25	S1		0.041+				0.0034+		
Y-393	38/8	S1		0.032+				0.012+		
B2-226	28/18	C		0.048+				0.049+		
G2-444	14/23	C/XII		0.031+	0.017+					
G1-250	37/11	C		0.041-	0.045-	0.011-			0.019-	0.0038-
B1-330	24/24	S3		0.0015-				0.0049-		
G3-387	20/22	S3		0.0048-				0.017-		
Y-97	34/12	S1		0.0064-				0.012-		0.043-
B1-141	28/21	S3		0.028-				0.036-		
B1-179	28/21	C		0.024-				0.034-		
B1-387	21/27	S3		0.045-				0.014-		
G1-130	37/11	C		0.043-		0.017-				
<b>G1-169</b>	8/37	S3/XII			0.0028-	0.032-				
G3-171	34/10	C			0.0055+	0.017+				
B1-101	38/10	C			0.031+	0.032+				
G1-157	22/27	S1			0.0082-					
B3-98	41/6	C			0.015+			0.031+		0.023+
G2-199	25/24	S3/VIII			0.018+			0.014+		
Y-247	37/9	C			0.021+			0.016+		
B3-111	15/31	S1			0.046+				0.014+	0.037+
B3-122	32/15	C/IX			0.039+				0.031+	
G3-338	26/18	C				0.0083+				
Y-120	39/7	C				0.0019+			0.017+	
G1-187	15/27	S1				0.049+			0.0086+	
B2-96	44/5	C				0.0082-				
B1-66	29/19	S1					0.007-			
B2-390	16/31	S3						0.0054+		
<b>B1-215</b>	35/10	S3						0.028+		0.041+
B1-133	40/9	C						0.034+		
<b>B3-215</b>	23/24	S3						0.047+	0.035+	0.017+
G3-358	16/29	S3						0.0018-		
B3-279	38/9	S3						0.0087-		
<b>G3-54</b>	32/14	C						0.011-	0.044-	
G3-206	38/8	S3						0.038-		0.045-
<b>B3-352</b>	16/31	S3							0.0091+	0.028+
B3-105	37/10	C							0.032+	0.0088+
B1-483	20/28	S3							0.0078-	0.0053-

<sup>a</sup> Markers highlighted in bold indicate that effects on the same traits were detected at least once at the 5% level in trial 1998

<sup>b</sup> Number of individuals with the marker presence versus absence,

<sup>c</sup> SY=starch yield, TSC=tuber starch content, TY=tuber yield, FE=foliage earliness, Gra=Gransebieth, Win=Windeby

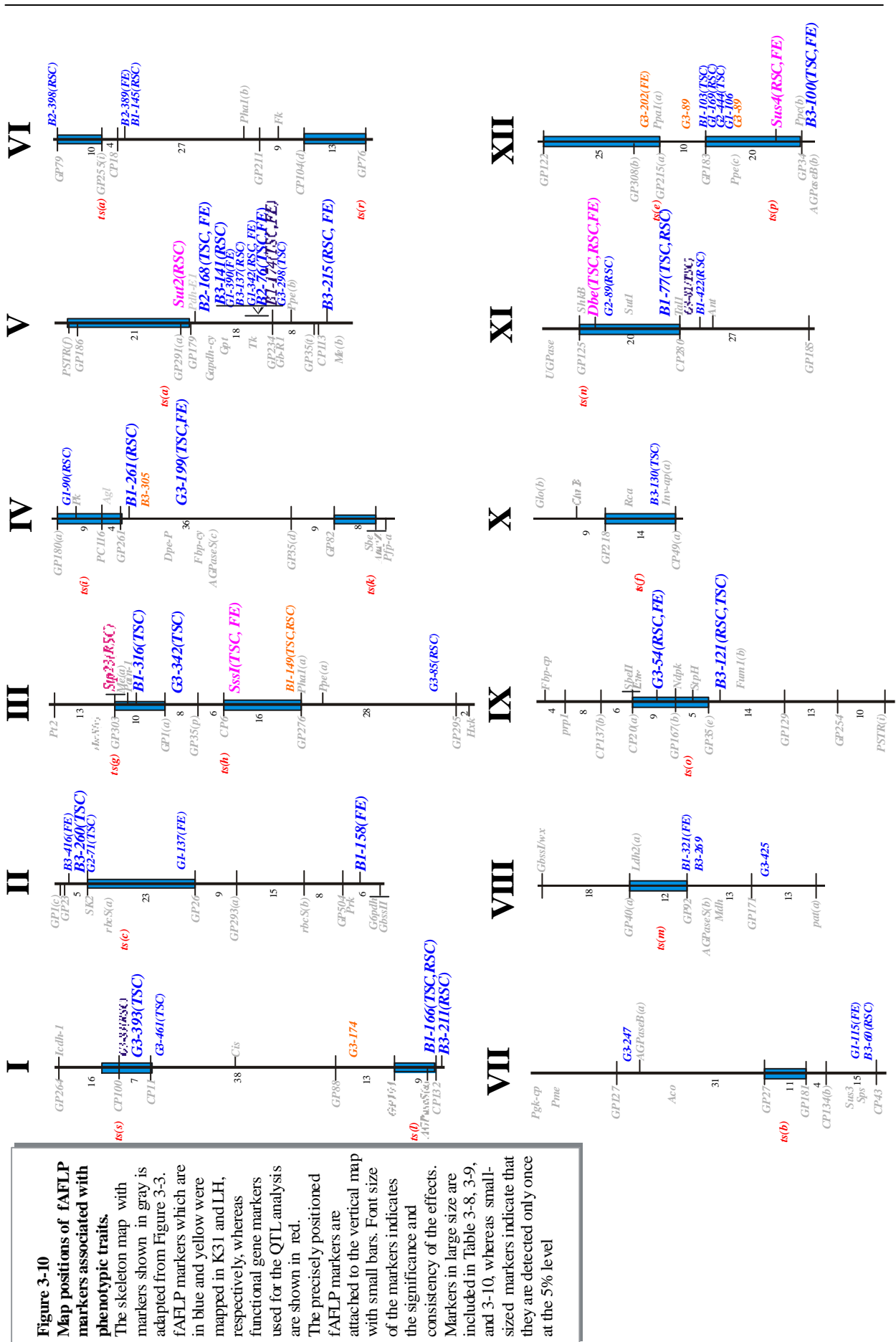
### 3.3.5 Assignment of map positions to QTL-linked fAFLP markers

For fAFLP markers to be useful in a QTL study, they are ideally required to be assembled into maps showing their chromosomal locations. For this purpose, a strategy was adopted which is based on the assumption that fragments of identical size occupy the same map position in different mapping populations (van Eck *et al.* 1995). The reference mapping population K31 and LH was genotyped using the fAFLP technique with the same restriction enzymes and the same primer combinations as used for the Z3 population. Polymorphic fragments generated with the same primer pairs and having the same size as fAFLP markers of interest in the tetraploid Z3 population were selected for linkage analysis. The genotypic data, alongside the previously generated framework of RFLP data were analysed for linkages using the MAPRF<sup>®</sup> program (see Sect. 2.11). Of the fAFLP fragments that had shown associations with phenotypic traits (at 5 % significance level), 50 common fragments were identified which segregated in the reference diploid mapping population K31, and could be mapped, therefore, to potato linkage groups (Figure 3-10, on page 83). Of these, 9 fAFLP fragments (G3-393, B1-166, B3-260, G2-71, B1-316, B3-121, B3-130, B1-77, B1-103) which showed associations with TSC in Z3, were mapped within or closely linked to intervals of QTL for TSC as detected in K31 on LG I, II, III, IV, IX, X, XI and XII. 20 fAFLP fragments, which were associated with RSC in Z3, were positioned on LG I, III, IV, VI, VII, IX, XI and XII. 16 fAFLP markers associated with FE were localised on LG II, III, IV, V, VI, VII, VIII, IX, XI and XII. These fAFLP markers were scattered throughout the 12 linkage groups but were, to some extent, clustered, as seen on LG I, II, IV, V, VI, VII, IX, XI and XII. In many cases, clustered markers tagged QTL for the same traits, such as marker clusters for TSC on LG I, II, V, XII, and clusters for RSC on LG I, IV, V, XI. The most remarkable cluster of the fragments is on chromosome V where 9 fAFLP marker loci were clustered in a region of less than 30 cM. Most of these markers (6) showed associations with FE (Figure 3-10).

### 3.3.6 QTL analysis through candidate gene strategy

Five functional genes (*Dbe*, *SssI*, *Stp23*, *Sut2* and *Sus4*) selected on the basis of known overlapping positions with QTL for TSC as identified in population K31 (Schäfer-Pregl *et al.* 1998) were selected and tested for associations with putative QTL in population Z3. Genotypic CAPS data, which were generated from individuals of Z3 at the 5 gene loci, was analysed for associations with TSC, RSC, and FE using the appropriate

(continue to page 84)



## Results

(continued from page 82)

tests as described in Sect. 3.3.3. Of the 5 genes tested, *Dbe-e* and *SssI*, located on chromosomes XI and III, respectively, showed association with TSC in population Z3 at the 1% level (Table 3-12). The effects were both from an S3 allele, the presence of the allele *Dbe-e* at locus *Dbe* resulted in increased TSC, while the allele at the *SssI* locus decreased TSC. The proportion of the phenotypic variance explained by the two loci was 3.5 % and 4.3 %, respectively. Besides the most significant allele *Dbe-e*, two other *Dbe* alleles from S3 were also detected as having significant effects (5 % level) on TSC at Gransebieth. At other gene loci no effects on TSC were detected in the tetraploid Z3 population. However, these loci showed associations with RSC. For instance, *Stp23* and *Sut2*, located on chromosomes III and V, respectively, were consistently detected to be associated with RSC at two locations at the 1 % level (Table 3-12). The alleles detected were from parent S3 at both loci and accounted for 3.7 % - 7 % of the phenotypic variance with increasing effects on RSC. No significant effects on FE were detected at these gene loci.

In addition, multiple alleles were identified at loci *Dbe*, *Sus4* and *SssI* which made possible the subdivision of the population at these loci into four or more genotypic classes. Comparison of the mean of median values of these genotypic classes was achieved by applying the Kruskal-Wallis H-test. Only *SssI* and *Dbe* showed significant effects ( $P < 0.05$ ) on starch and RSC respectively at Windeby (data not shown).

**Table 3-12 Functional gene markers showing associations with multiple phenotypic traits**

fAFLP marker	LG <sup>a</sup>	No -/+ <sup>b</sup>	$\chi^2$ -test <sup>c</sup>	Ori. <sup>d</sup>	Tuber starch content				Reducing sugar content			
					Gransebieth		Windeby		Gransebieth		Windeby	
					$P^e$	$R^{2f}$ (%)	$P^e$	$R^{2f}$ (%)	$P^e$	$R^{2f}$ (%)	$P^e$	$R^{2f}$ (%)
<i>Stp23</i>	III	22/102	1:5	S3					0.008+	4.7	0.002+	7.0
<i>LeSUT2</i>	V	63/89		S3					0.017+	3.9	0.012+	3.7
<i>Dbe-a</i>	XI	82/68	1:1	S1							0.014-	4.7
<i>Dbe-b</i>	XI	101/97	1:1	S1								
<i>Dbe-c</i>	XI	129/58		S3	0.02+	2.5						
<i>Dbe-d</i>	XI	133/52		S3	0.04+	2.0						
<i>Dbe-e</i>	XI	143/52		S3	0.008+	3.5						
<i>Sus4</i> *	XII	78/66	1:1	S1					0.06-	2.3	0.038-	2.6
<i>SssI</i> *	III	93/98	1:1	S3			0.004-	4.3				

<sup>a,b,c,d,e,f</sup> see Table 3-10 (on page 77-78) for references

\* two alleles of genes *Sus4* and *SssI* were identified, only the ones showing significant associations with RSC and TSC in the t-test were included in the table.

### 3.3.7 Cloning of fAFLP fragments

For breeding-oriented QTL analysis, it is highly desirable that significantly associated markers are amenable to marker-assisted selection. As a first step towards this purpose, fAFLP markers associated with a trait of interest need to be cloned. Since recovery of fAFLP DNA fragments directly from the automated fluorescent genotyping system is not technically accessible to date, this was bypassed by the use of the radioactive AFLP approach. Based on the results of association analysis, the five most significant and reproducible fAFLP fragments B3-117, 121, 215, 409, and 425 (see Tables 3-8, 3-9, 3-10, on page 73, 75-76, 77-78) generated with primer combination E-act/M-cta (B3) were selected for cloning. 10 individuals of Z3, for which the target fragments were well separated from the surrounding bands based on fAFLP fingerprinting, were tentatively selected for cloning. The same preamplification products of the two parents and ten selected individuals previously used for fAFLP were used directly for selective amplification. AFLP procedures were followed as described by Meksem *et al.* (1995). The AFLP size marker 30-330 bp with a 10 bp-ladder between neighbouring fragments was used to identify the fragments of interest. Of the five fragments examined, two fragments B3-409 and B3-215, at the expected position on an autoradiogram, gave rise to an identical segregation pattern to that of the fAFLP fragments. The other three fragments, which appeared faint on the autoradiogram, were not pursued further. The two identified bands were excised from the polyacrylamide gel, eluted and re-amplified with the corresponding non-radiolabelled primer pair to recover the DNA. This was successfully achieved for fragment B3-215 but failed for B3-409. The re-amplified PCR product of B3-215 fragment was subsequently cloned into the pGEM<sup>®</sup>-T vector. Putative recombinant colonies were either confirmed by colony PCR which was performed directly using bacterial colonies on agar plates as templates or alternatively, by digestion of the extracted plasmid DNA with restriction enzymes. Plasmid DNA from 10 confirmed colonies was prepared and subjected to sequencing. All 10 sequences showed the selective primer sequence of *EcoRI* at one end and the *MseI* primer sequence at the other. The analysis of the 10 sequences revealed three different sequence classes. One was 218 bp in length; two were 219 bp sequences; the third one was 217 bp in size with 7 members, which showed high similarity (98-99%) and differed from one another only by single nucleotides (Figure 3-11). The majority class had a 2 bp discrepancy from the expected 215 bp. This discrepancy may have resulted from the automated sizing by the GeneScan<sup>®</sup> system by which size assignment may have 1-2 bp variations depending on the size categories defined. Taking into account

Primer E-cta

5' - GATGAGTCCTGAGTAActaGGTTCG (G/A) CAAACTTTTGTTTTCTTTTGCCTA  
 CTCCTCGGACACAACAACCAATATTTATTTCACTC (C/T) AAAAGA (G/T) TAC  
 AAGTGAAATACTACAAGAGAGAAAGAAGATCAAATGCCTTTGAAGAT (G/A) A  
 (G/A) AAGGCAAG (T/C) GAGAGGTGTGT (T/C) ACAAATGAATTAGGAACTAC  
 CTATTTAT (A/G/T) GGagtGAATTGGTACGCAGTC - 3'

Primer M-act

**Figure 311** The combined consensus sequence from the 7 sequences of the cloned fAFLP fragment B3-215. Nucleotide variations among the 7 sequences were indicated in parentheses. Primer sequences were underlined and lower case residues were selective nucleotides.

this fact, the 7 sequences were assumed to be the correct ones. A BLAST search for homologous sequences in the data bases indicated that this sequence showed 88% homology at the nucleotide level to a downstream sequence segment of the ORF of the gene coding for chlorophyll a/b binding protein (*Lhcb1*, accession No. AB012636) in *Nicotiana sylvestris*. For confirmation and PCR-marker conversion of this cloned fragment, various approaches were tested. First, since the complete sequence was known, a primer pair was designed by adding four more sequence-specific nucleotides at the 3' end of each selective primer. The idea was that using more selective primer to amplify the pre-amplification product used for fAFLP generation would result in a single or few fragments resolvable on agarose gel with segregation pattern identical to the fAFLP B3-215 marker. Although a single fragment with the expected size was obtained under stringent amplification condition, no polymorphisms were exhibited between Z3 individuals. This may be due to the fact that the selectivity of the primers was reduced or even lost after addition of the four sequence specific nucleotides at the 3' end of the primers. An alternative option for confirmation could be DNA blotting, by which gel-electrophoresis-resolved AFLP fragments are transferred on to membrane followed by hybridization with a probe from the cloned AFLP fragment. This approach was, however, not applied in considering that the final goal was to convert the fAFLP marker into PCR-based marker, which, when achieved, would also lead to the confirmation of the segregation pattern. For marker conversion, more sequence information is needed. The high sequence similarity of the cloned AFLP fragment to the gene *Lhcb1* of *Nicotiana sylvestris* was used to



## Results

achieve the goal by amplification of the homologous gene segment in potato population Z3. A primer pair was designed with one primer being derived from the most conserved region of *Lhcb1* gene, the other from the sequence of the cloned fragment. The two primers encompassed the *MseI* restriction site of the AFLP fragment. A single fragment of 500bp in size was amplified from both parents S1 and S3 using genomic DNA as template. This PCR product was digested with restriction enzyme *MseI* to test if the *MseI* site was the cause of yielding the fAFLP fragment. However, the cleavage of the PCR product with *MseI* did not detect CAPS between the two parents. Sequencing of the PCR fragment from S1 revealed high sequence homology (93 %) to the cloned fragment B3-215 and also an *MseI* site as expected (Figure 3-12). The point mutations between S1 and S3 allowed designing allele-specific primers, however, PCR amplification failed to distinguish them. This is not unexpected since the allele-specific primers were base on the difference of only one allele from S1 whereas the other three remained unknown. Consequently, the fAFLP fragments B3-215, while successfully cloned, is still to be confirmed.

1	<u>GAGAACCTTG</u>	<u>CTGACCACCT</u>	<u>TGCTGAAGTT</u>	TGTGTCTTGC	GATTAGCAAT
51	CGTAGGGATT	TTCTTATAAC	AGGATCACAT	GATCGTTCCA	TACGACGTTG
101	GGATCGTACT	GATGATGAAG	CTCTTTTAT	TGAGGTGTTA	TATTTTGCAT
151	CAGTTTTAGA	CAGTTAAAAA	GAATAATTTA	TAAAGTTGCT	ATTTGGAATT
201	GCAGGAAGAG	AAAGAAAAGA	GATTGGAAGA	GATGTTTGAA	TAAGATATCG
251	ATAATGTATT	TGAGAATAAG	TATGTGTTGT	TGAAGCCGAA	TATATAGAGA
301	GTAATGGAAT	CGCAGCTGCT	ATATCTAAAG	GTAGCTAATA	AATAGTAAAT
351	GAGATAACAA	TAAAAAGAAC	ACCATGAATT	<b>AACGAGGTT</b>	<i>GGCAAAATTT</i>
401	<i>TGATTTCTTT</i>	<i>TGCCTATTCC</i>	<i>TCGGACACAA</i>	<i>CCAACCAATA</i>	<i>TTTTTACCAT</i>
451	<i>TCCAAAAGAG</i>	<i>TACTAGTAAA</i>	<i>ATACTACAAG</i>	<i>AGAGAAAGAA</i>	<i>GATCAAATGC</i>
501	<i>CTTAGAAGAT</i>	<i>GAGAAGGCAA</i>	<u>GTGAGAGGTG</u>	<u>TGTTAC</u>	

**Figure 3-12 Extended nucleotide sequence of the cloned fAFLP fragment B3-215.** Sequence was derived from parent S1, nucleotide residues in italics indicate the sequence of the cloned fAFLP fragment, others are extended sequence starting from *MseI* site (highlighted in bold). Primers sequences are underlined.

## 4. DISCUSSION

---

### 4.1 Mapping functional genes involved in carbohydrate metabolism and transport

#### 4.1.1 STS as an approach for functional gene mapping

In the past decade, many advances have been made in the construction of molecular maps for crop plants using various molecular marker tools such as RFLP, RAPD, SSR and AFLP. These types of molecular linkage maps play an important role in the genetic analysis of agronomic traits including map-based gene cloning and QTL analysis. However, all these types of markers are largely anonymous and of unknown function, limiting further exploration of the genetic basis underlying the marker-tagged agronomic traits studied. Thus, the construction of linkage maps with markers tagging known functional genes is highly desirable and, when achieved, could take genetic mapping to a further level. Typically, Southern hybridization of a segregating population with a cloned gene as probe is the common way to localize the gene on molecular linkage map and 22 genes have been mapped through this approach (Gebhardt *et al.* 1994). Although RFLP is notably reliable for mapping, the application is limited by the availability of a clone for the gene of interest and additionally, the marker created could hardly be incorporated into MAS for screening plants on a large-scale because the technique is laborious and time-consuming.

In the present study, public databases were systematically searched for sequences of plant genes with a functional role in carbohydrate metabolism and transport. Sequence information was used to develop PCR-based marker assays that allowed localization of corresponding potato genes on existing RFLP linkage maps. This approach takes advantage of existing sequence resources in the databases for gene-specific markers, thus enabling the efficient and successful construction of molecular maps for functional genes with minimal additional requirements. The highly polymorphic mapping populations K31 and LH offer a good foundation for the positioning of functional genes on potato linkage maps through an STS approach. All of the genes in this investigation were mapped through this approach using one or other of the two populations except gene *Tk* for which no polymorphism was detected. Alternatively, *Tk* was mapped by a RFLP assay using a PCR product of *Tk* as a probe. Using

## Discussion

---

the STS approach in combination with RFLP, an initial potato molecular function map based on RFLP mapping of cloned genes (Gebhardt *et al.* 1994) was expanded about four-fold.

To map functional genes efficiently and successfully through the STS approach, selection of the appropriate gene regions for primer design is critical since the majority of gene sequences available in databases are cDNA sequences. It was found to be important for primer design to avoid the exon-intron boundaries and to position the two primers at a distance of 500-800bp apart from each other on the cDNA sequence so that PCR products of sufficient length encompassing introns could be amplified for identifying polymorphisms. Two methods gave an indication of such exon-intron junctions. First, some sequence motifs such as CAGG or AAGG may be a hint of the junction sites (Lefebvre and Gellatly 1997). Thus primer design should avoid these sequence motifs. Second, if the genomic gene sequence of another related species is available, the comparison of the cDNA sequence with the genomic sequence could lead to the identification of putative introns.

### 4.1.2 Validity of the STS approach for mapping functional genes

In the present study, the first technical choice for function mapping was the STS approach because of (i) the availability of large number of nucleotide sequence information of genes in the GenBank and EMBL databases; (ii) the ability to specifically detect by PCR genes known in different genomes using degenerate primers complementary to exon sequences (Larrick *et al.* 1989, Holland 1993, Gould *et al.* 1989, McPherson *et al.* 1991 and Van Campenhout *et al.* 2000); (iii) the possibility of directly sequencing PCR products, which is the best way to validate gene homology and (iv) the feasibility of routine application of the markers in MAS. The reliability of the STS approach was confirmed by sequencing the PCR products and comparison of their map positions. First, the specificity of the PCR amplification was verified by sequencing three PCR products generated with primers from potato sequence and all six PCR products obtained with degenerate primers from other species than potato. The former cases showed 95 % -100 % similarity and the latter cases showed 80 % - 95 % similarity with the reference sequences. Second, seven genes were mapped by both the STS approach and by RFLP using different mapping populations and all were localized in similar positions on potato linkage map. This suggests that the employment of an STS approach for function mapping is reliable.

### 4.2 The potato function map

#### 4.2.1 PCR-based markers for genetic analysis and marker assisted selection

The functional genes mapped by the STS approach in the present study, together with other previously mapped genes, creates a framework of 66 gene markers opening up new possibilities for various genetic studies in potato. On one hand, since the genes mapped are all genes with known function and mostly involved in carbohydrate metabolism and transport, thus candidate gene strategy becomes accessible to analyzing QTL for tuber starch and sugar content in different populations. This could be achieved by genotyping with the candidate gene markers for their segregation in populations considered and testing their significance of effects on the traits of interest. This approach can eliminate the generation of the genome-wide markers by anonymous marker technology, and has been proven fairly efficient in QTL studies as described in Section 3.3.6. On the other hand, since the gene markers as defined in this study are PCR based, these markers, together with previously developed sequence based anchor markers (Oberhagemann *et al.* 1999) are a valuable tool for genetic analysis and marker-assisted selection in potato. They (i) cover a considerable proportion of the potato genome, (ii) are available as locus specific sequence information and (iii) can be analyzed by PCR with small amounts of genomic DNA as template; (iv) may be further optimised for large-scale screening through the use of high throughput techniques like molecular beacons (Tyagi and Kramer 1995; Tyagi *et al.*, 1998), thereby eliminating gel electrophoresis. Marker assays will have to be verified and adapted in order to be informative, however, when applied to germplasm other than the material used for gene mapping in this study.

#### 4.2.2 Molecular function maps for comparative studies of plant genomes

Most of the genes for which map positions are reported in this paper are housekeeping genes functioning in general plant metabolism. Even distantly related plant species have in common these genes with similar functions and a high degree of sequence conservation. In different plant species, molecular maps based on common genes may be used not only for comparative studies of plant genome evolution, but also for functional comparisons. Few molecular maps are, however, currently available in plants for specific sets of genes of known biochemical and physiological effect. Similar studies have been carried out in maize and sugar beet (Causse *et al.* 1995, Schneider *et al.* 1999). The sequencing of whole plant genomes as done

## Discussion

---

for *Arabidopsis* (Bevan *et al.* 1998, Lin *et al.* 1999) or the construction of linkage maps based on expressed sequence tags as done in rice (Kurata *et al.* 1994) and maize (Chao *et al.* 1994) makes available information on genomic positions of functionally and structurally similar genes which can also be used for genomic comparisons. For the potato and tomato genomes, extensive colinearity has been demonstrated (Bonierbale *et al.* 1988). In tomato, a genetic map including isozyme loci and more than 300 randomly selected EST markers has been constructed (Tanksley *et al.* 1992). Based on their DNA sequence, putative functions have been assigned to tomato EST markers (Ganal *et al.* 1998). The availability of a potato function map made possible the evaluation of synteny between potato genes and tomato EST markers, to which the same functions have been assigned based on sequence similarity (Ganal *et al.* 1998, Tanksley *et al.* 1992). CAPS markers for potato genes *Icdh-1* and *Aco* mapped to positions on linkage groups I and VII, respectively, which are in good agreement with isozyme loci *Idh-1* and *Aco-2* on the tomato molecular map. Potato genes *UGPase* on linkage group XI, *Pgk<sub>cp</sub>* and *Pha2*, both on linkage group VII, occupy similar map positions as tomato markers *CT182*, *CT114* and *CD54*, respectively, which code for the same enzymes (Ganal *et al.* 1998, Tanksley *et al.* 1992). In four cases, the potato gene marker identified a single locus, whereas two or more loci were identified by the corresponding tomato EST marker. The four loci were: *AmyZ* (LG IV), *Stp1*, *Tk* (both LG V) and *Ndpk* (LG IX), which correspond to tomato marker loci *CT224B*, *CD31A*, *CD38B* and *CT225A*, respectively, to which the same function has been assigned based on sequence similarity (Ganal *et al.* 1998, Tanksley *et al.* 1992). In one instance, the CAPS marker for the potato *Ppe* gene identified three loci in potato on linkage groups III, V and XII, whereas tomato marker *CT248* for *Ppe* detected only the locus on linkage group III. No correspondence between map positions was found for the potato *Me* gene for NAD<sup>+</sup>-dependent malic enzyme (LG III and V) and tomato marker *CT201* which is similar to NADP<sup>+</sup>-dependent malic enzyme. The DNA sequences of the two markers share only 56 % similarity. The map positions of potato genes for fructose-1,6-bisphosphatase (*Fbp<sub>cp</sub>* on LG IX and *Fbp<sub>cy</sub>* on LG IV) were both different from the position of tomato marker *CD5* (LG X) which has sequence similarity with the same enzyme. The potato homologue of a tomato glycolate oxidase gene (*Glo*), although 96.6 % identical with the tomato sequence from which it was derived (Table 3-2), detected two loci on potato linkage groups VII and X, which were in disagreement with the position on tomato linkage group II of tomato marker *CD79* which is identical to the tomato *Glo* gene sequence used for primer design (M. Ganal, personal communication).

In potato and tomato, orthologous genes with conserved sequences and the same function assignment are expected to occupy similar map positions. This was, in fact, observed for most of the genes examined, when taking also into account the different numbers of loci detected in several cases in the two species. The difference in the number of loci identified with the same gene marker might be due to (i) the fact that not all loci of a gene are identified due to the limited DNA variation in a given mapping population or limited power of the techniques used for mapping. For this reason, the number of loci detected may be considered as a minimum estimate. Consequently a locus may be detected in both species, another locus may be identified in one species but not in the other or vice versa; (ii) the nature of the marker assays used for mapping. The RFLP assay identifies, depending on the hybridization stringency used, loci with 80 % or less sequence identity with the probe, whereas a PCR-based assay requires higher identity between primer and target sequence and is, therefore, more gene specific. Alternatively, the potato and tomato genomes may have indeed different numbers of copies of certain genes (Gebhardt *et al.* 1991). Three gene markers, however, did not occupy syntenic positions on the potato / tomato maps. First, this might be due to erroneous functional assignment when sequence similarity is not extremely high as in the case of malic enzyme (*Me*) where the potato gene shared only 56 % similarity with the tomato EST. Second, in the case of gene families with members that are not linked, different loci might be polymorphic in the two species compared. Third, members of gene families with high interspecific sequence conservation compared to other members of the same family may not always be truly orthologous. Differential selection pressure acting on members of a gene family after speciation may have resulted in higher sequence divergence between orthologous members than between paralogous members. This might explain why a glycolate oxidase (*Glo*) gene marker, although highly conserved between potato and tomato, was located on different linkage groups in the two genomes.

### 4.2.3 Candidate genes for TSC (TSC)

Quantitative trait loci (QTL) for TSC and tuber yield have been mapped in the populations K31 and LH (Schäfer-Pregl *et al.* 1998). The marker intervals with approximate positions of tuber starch QTL as previously identified are indicated in Figure 3-3. The same populations were used for PCR-based mapping of genes involved in carbohydrate metabolism and transport. Comparison of the QTL map with the molecular function map reveals a number of correlations between map positions of QTL for TSC and functionally related loci. This was

## Discussion

---

expected, due to the map coverage by both the QTL and the functionally related loci. Overlapping chromosomal positions are observed, either when the gene(s) responsible for a QTL effect are linked but functionally unrelated to a marker gene, or when the marker gene itself is responsible for the QTL effect. Evaluation of the latter possibility has to be knowledge-based, considering the physiological, biochemical and molecular role of candidate genes.

The most direct candidates controlling TSC are genes encoding starch-metabolizing enzymes. Sixteen genes of this class were mapped yielding eighteen loci on the potato molecular function map (Gebhardt *et al.* 1994 and this study). ADP-glucose pyrophosphorylase is a key enzyme of starch biosynthesis (reviewed in Frommer and Sonnewald 1995). Markers for the two subunits S and B of potato AGPase identified five loci one of which, *AGPaseS(a)* on linkage group I, was positioned in the same 9 cM marker interval as the QTL *ts(l)* (Figure 3-3) having a small effect on TSC (Schäfer-Pregl *et al.* 1998). One of four mapped starch synthase genes, *SssI* on linkage group III, is linked to QTL *ts(h)*. Genes *SbeI* for starch branching enzyme I and *Dbe* for debranching enzyme are linked to QTL *ts(k)* and *ts(n)* on linkage groups IV and XI, respectively. Five genes coding for starch degrading enzymes were mapped: *Stp23* on linkage group III, *AmyZ* and *AgI* both on linkage group IV, *Stp1* on linkage group V and *StpH* on linkage group IX. All five loci are linked to QTL for TSC (Figure 3-3). The clearest positional correlation was between *Stp23* and QTL *ts(g)* on linkage group III because both the QTL and the marker gene are most closely linked to the same anchor RFLP marker *GP303* (Figure 3-3 and Schäfer-Pregl *et al.* 1998). This suggests that natural allelic variants of genes controlling starch synthesis and degradation may contribute to quantitative effects on TSC.

Accumulation of tuber starch also depends on availability of substrate for starch synthesis and, therefore, more indirectly on photosynthesis, sugar metabolism, transport and energy supply (reviewed in Frommer and Sonnewald 1995). Allelic variants of genes operating in these metabolic pathways may be responsible for variation of TSC. Known genes in these pathways were placed, therefore, onto the potato molecular function map. The list of the genes mapped is, by no means, complete. QTL for TSC were linked, among others, to loci for ribulose biphosphate carboxylase, small subunit (*rbcS-c* and *ts(c)* on LG II, *rbcS-1* and *ts(g)* on LG III), ribulose biphosphate carboxylase activase (*Rca* and *ts(f)* on LG X), sucrose synthase 4 (*Sus4* and *ts(p)* on LG XII), pyruvate kinase (*Pk* and *ts(i)* on LG IV), pyrophosphate fructose-6-phosphate 1-phosphotransferase,  $\alpha$  subunit (*Pfp-a* and *ts(k)* on LG

## Discussion

---

IV), malic enzyme (*Me(a)* and *ts(g)* on LG III), plasma membrane H<sup>+</sup>-ATPase 2 (*Pha2* and *ts(b)* on LG VII) and soluble inorganic pyrophosphatase (*Ppal(a)* and *ts(e)* on LG XII), soluble acid invertase (*Pain-1* and *ts(g)* on LG III) and apoplastic invertase (*Inv-ap(a)* and *ts(f)* on LG X)

In many cases, the same QTL was linked to several candidate gene loci. For example, at least three loci, *rbcS-1*, *Stp23* and *Me(a)*, overlap with QTL *ts(g)* on linkage group III and three others, *SbeI*, *AmyZ* and *Pfp-a*, overlap with QTL *ts(k)* on linkage group IV. But from this study it is unclear whether all are contributing equally or if only one is the primary contributor to the overlapping QTL or if other unknown genes at these regions are responsible for the effects. Several strategies can be adopted for further study to validate a putative candidate gene. First, QTL analysis may be carried out based on activity measurements of the candidate enzymes as performed in maize (Causse *et al.* 1995); second, transformation of a genotype with various alleles may be conducted for expression analysis and functional analysis of molecular variants. In this case, however, allele-trait association may go undetected if different alleles have a similar influence on the trait.

Three of the seventeen QTL for TSC (*ts(s)*, *ts(a)* and *ts(r)*) shown in Figure 3-3 were not linked to any known candidate gene locus. QTL *ts(s)* on linkage group I was linked to acidic and basic  $\beta$ -1,3-glucanase genes (*GluA* and *GluB*) which, like chitinases (*ChitB* on LG X), are related to pathogenesis and defense (Kombrink and Somssich 1997) response to diseases rather than to starch accumulation in tubers. Therefore, they are not considered as candidate genes for this trait despite their linkage to QTL for TSC.

The present study focused mainly on the structural genes coding for various enzymes functioning in the carbohydrate metabolism and transport. It is clear that regulatory genes controlling transcription, translation and posttranslational modification of the starch metabolic enzymes play also a pivotal role in these processes. For example, the *p1* locus in maize which encodes a transcription factor regulating the synthesis of compounds in the flavonoid pathway was proved to be an important regulatory locus as a QTL for resistance of maize to corn earworm (McMullen *et al.* 1998). However, regulatory factors have not been reported for starch metabolism to date. While the physiology, biochemistry and molecular biology of carbohydrate metabolism and transport in plants have been extensively studied, the information on the total number of genes relevant to TSC is incomplete and will remain so for the foreseeable future. Nevertheless, there are a few lines of evidence suggesting that a



molecular function map for the genes currently known to control carbohydrate metabolism and transport is a good basis for explaining at the molecular level QTL for starch- and sugar-related agronomic traits in plants. Apart from the results of our study as outlined above, in maize, all mutations affecting the starch content of the kernel that have been identified at the molecular level affect enzymes of carbohydrate metabolism (Neuffer *et al.* 1997). QTL for sugar content of kernels and other traits of maize are linked, for example, to QTL for enzymatic activity of sucrose phosphate synthase (Causse *et al.* 1995). Moreover, a first QTL for sugar content of tomato fruits has been characterized on the molecular level and shown to be controlled by molecular variants of an apoplastic invertase gene (Fridman *et al.* 2000).

### 4.3 Single nucleotide polymorphism (SNP)

SNP is a new type of molecular marker that is very useful and robust in genetic association analysis. SNPs are detected by various techniques, for instance, by gel-based sequencing (Nickerson *et al.* 1997), high-density variation-detection DNA chips (VDAs) (Wang *et al.* 1998), and direct approach based on public databases (Picoult-Newberg *et al.* 1999). In the present study, an attempt was made to identify the potential SNPs existing in the gene *Stp23* by sequencing PCR products of the gene derived from different genotypes. The template DNA used for sequencing was PCR product amplified with the same primers from both diploid and tetraploid potato lines. The objective was to investigate the feasibility of detecting potential SNPs by sequencing PCR products of the gene of interest and, to characterize the SNP distribution in functional genes. The results of SNP obtained, although they need to be confirmed, represent a promising beginning for further identification and application in potato genetic studies.

In this study, the identification of SNPs has been demonstrated (Sect. 3.2) by visually examining the electropherogram of sequence data from a PCR product. Overlapping peaks at a single nucleotide position indicates a potential SNP (Figure 3-6). The SNPs identified by this way reflected the heterozygosity of the genotype. When comparing the peak height of a nucleotide signal in the homozygous state with those in the heterozygous state (potential bi-allelic SNPs), a characteristic drop in peak height at heterozygous positions for two peaks is observed on the electropherogram and the two peaks result frequently in a base called 'N' (Figure 3-6). This feature can be used to detect heterozygous positions (SNPs) from sequence data. The basis of this feature may be the different dosage of the single nucleotide templates.

## Discussion

---

In effect, when sequence data were based on a PCR product which is the mixture of two alleles (in diploids), nucleotide dosage of the template can be reflected by the signal intensity or peak height on the electropherogram. Obviously at heterozygous positions, the template of each nucleotide is only half concentrated compared to those at homozygous positions. Thus, a drop of peak height at SNP positions is expected on the electropherogram. This feature was observed for all bi-allelic SNPs identified with only exception of the consecutive tri-allelic and bi-allelic SNPs found in tetraploid line Taiga (Figure 3-6).

The only drawback for the identification of SNPs through direct sequencing of PCR products is that it requires high quality sequence data to give a reliable discrimination between heterogeneity of alleles and signal noise. High quality sequence data is needed for confirmation of putative heterozygous / homozygous position.

Sequencing from cloned PCR products gives high sequence quality, however, additional work like cloning of PCR products into vectors and multiple clones are required. For this reason, sequencing directly from PCR products has more advantages in terms of efficiency, particularly for tetraploid lines.

Besides the identification of SNPs as heterozygous nucleotides within single lines, sequence assembly of different lines can detect mismatches for potential SNPs among different lines. Such SNPs reflect the frequency that a position within the assembly has heterogeneity in nucleotide composition. This depends on the number of sequences compared and the genetic diversity among them.

For large-scale identification of SNP by sequencing, efficiency could be enhanced by sequencing pooled DNA of different sources rather than numerous individuals as done in the present experiment. Additionally, the procedures of sequence analysis could be automated by use of computer programs *Phred* for base calling, *Phrap* for sequence assembly and *Consed* for sequence assembly editing (Nickerson *et al.* 1997).

In general, a potential SNP for a single copy gene may be considered as real when the fluorescence-based sequence was of high quality. However, care must be taken when interpreting the results. First, the base-calling error of sequencing exists, which can be overcome by repeated sequencing of the same sequence segment in different genotypes. Second, the possible paralogous of a gene family should be considered. In cases when the gene examined is a member of a multigene family, the candidate sites that appear

## Discussion

---

heterozygous in a sequence, may be the result of mismatches between paralogous members of the multigene family. This issue can be further addressed by assessing the genetic distance between members of the sequences.

For aspects of application, SNP data of genes are required, which are representative for germplasm potentially used in breeding, so that SNPs identified are likely to be informative in any breeding families, thus increasing the potential of applicability of SNP data. Therefore, a survey of SNP using a sufficient number of materials is needed. The sequences assembled for SNP discovery in gene *Stp23* represented some commercial and elite breeding lines used in breeding program. They may not be sufficient to achieve a good coverage of commercial breeding materials. They may also represent only a limited genetic base of potato since the cultivated potato in Europe was derived from a limited number of *Solanum tuberosum* ssp. *andigena* (Bradshaw *et al.* 1994). Therefore sequence differences observed in the 14 lines represent only part of the SNPs potentially existing in the potato gene pool.

In the SNP analysis of gene *Stp23*, eight exonic SNPs were revealed, of which four derived from the T710 parent of population LH and one from the P3 parent of population K31. These SNPs resulted in amino acid changes, which may further modify the gene function and expression, and ultimately affect the phenotypic trait like TSC. Interestingly, a remarkable effect on TSC was observed at the locus *Stp23* in both two populations K31 and LH (Schäfer-Pregl *et al.* 1998). This coincidence may imply that *Stp23* could be a causal gene affecting the TSC and the non-synonymous changes of amino acid observed in this gene could be the cause. It can not be ruled out, however, that SNPs at other regions, for example in the promoter region, could also be the cause for the phenotypic variation of TSC. When the gene causing the effect is validated, the effect of the allelic variations of the genes could be actually the real genetic factors affecting the quantitative trait, thus circumventing the map-based cloning of the QTL.

The SNPs identified occurred in exons as well as in introns. The former ones were much less frequent than the latter ones. This feature was also observed in other organisms (Li and Sadler 1991). In addition, SNPs are frequently distributed in the total genome of all organisms examined so far. For instance, a frequency as high as 1 in every 202 bp has been observed in mouse genome (Lindblad-Toh *et al.* 2000). Because of this, SNPs have many advantages and much potential for application. First, it can facilitate the development of high-density genetic maps because of the high occurrence in the genome. Second, since the

detection of SNPs can be focused on structural genes which are more valuable than anonymous regions. Therefore, SNP, as a new marker system, can provide an alternative to the STS approach for function mapping. Third, this approach could identify a number of sequence variants in exons or promoter regions that may lead to the alteration of gene function or gene expression. This could be associated with phenotypic variation in the population. The allelic variations of the gene may actually be the genetic factors affecting the quantitative trait. Finally, as the sequence context of the SNPs is already known, it is possible to exploit some of the automated approaches available to genotype a population on a large scale.

## 4.4 QTL analysis

### 4.4.1 Strategy of QTL-tagging

In potato, QTL analysis is largely constrained to the experimental diploid populations to avoid the complexities of tetrasomic inheritance (Bonierbale *et al.* 1994; Leonards-Schippers *et al.* 1994; Van Eck *et al.* 1994, Freyre *et al.* 1994, Van den Berg *et al.* 1996a, b; Douches and Freyre 1994; Freyre and Douches 1994; Schäfer-Pregl *et al.* 1998). Studies performed on diploids provide informative insight into the genetic nature of the trait studied. Making inferences from the diploid to the tetraploid level is not straightforward, however, since genes affecting quantitative traits might have different effects at the diploid and tetraploid ploidy levels (Groose *et al.* 1988). Thus, it is highly desirable that breeding-oriented QTL analysis is performed on tetraploids.

In general, QTL analysis requires construction of a linkage map with markers evenly distributed over the genome using the population from which phenotypic data were collected. However, due to the complexity of tetraploid genome structure, linkage construction in tetraploid potato is difficult. First, a total of 96 linkage groups (four homologous groups of 12 chromosomes for each parent) potentially exist for a tetraploid population of potato. For linkage mapping, an extremely large number of markers, preferably codominant have to be generated to have a sufficient number of markers on each group. Second, bridging markers need to be generated to identify the homologous groups. Apparently, the efforts leading to the linkage construction is expensive and time consuming, which has been demonstrated by Meyer *et al.* (1998). In their experiment, 573 AFLP markers that segregate in three segregation types (1:1, 1:3, 1:5) were selected for linkage analysis from the fragments

## Discussion

---

generated with 39 AFLP primer combinations. These markers could be organized into 30 and 26 linkage groups for the maternal and paternal maps, respectively, representing 25 % coverage of the genome. Only 8 homologous groups were identified because of the limited number of bridging AFLP markers linking homologous groups.

In present study, without a prior construction of linkage map, QTL were tagged by fAFLPs using fluorescent sequencing system in combination with t-test regression analysis and U-test. This is the first study on QTL tagging in a tetraploid crop plant using fAFLP marker technology. This strategy was adopted because, first, fAFLP is a semi-automated genotyping technique which allowed to generate with high efficiency sufficient amount of molecular markers in a short period of time compared to other available marker technologies. Second, the dominant nature of fAFLP markers being scored as presence versus absence, meets the requirements for QTL tagging by parametric t-test or non-parametric U-test and the quantitative fAFLP data also allowed correlating fAFLP markers with phenotypic traits by regression analysis. The power of these statistic tools may not be equally high when applied to data of markers with different segregating types. For example, segregation ratio deviated from 1:1 may reduce the power of the t-test. However, these two methods, when combined, can complement to each other and could give a fairly good association test for markers with different segregation types. And finally, this strategy needs not to generate so many markers as required for linkage mapping in tetraploid. In effect, when sufficient markers are generated to cover 12 chromosomal linkage groups irrespective of the homologous groups, further increase in number of markers may enhance the precise detection of QTL but may not significantly lead to the identification of new QTL. In the present study, more than 200 polymorphic fAFLP markers were estimated to be generated with 7 primer combinations (see Table 3-5). These markers, although much less than sufficient to generate linkage groups or even to identify marker pairs linked in coupling-phase, may give a fairly good overall chromosomal coverage. Van Eck *et al.* (1995) generated a total of 264 markers with six AFLP primer combinations and demonstrated a good overall genome coverage of the diploid potato genome. The fAFLP markers generated in this experiment, therefore, have the potential to tag the QTL dispersed in the total genome of Z3.

This approach, while successful in QTL-tagging, suffers some inherent weakness. On one hand, despite the fact that a number of fAFLPs were inferred from common AFLP markers identified in diploid reference mapping population (see Sect. 3.3.5), most fAFLP markers of interest still remain unknown in terms of their chromosomal positions. Consequently, it

## Discussion

---

remains unclear (i) whether QTL markers which showed associations with the same trait were linked or independent, (ii) how many QTL were identified and, (iii) to what extent the QTL can be exploited for MAS in the population, which, in case when the number of QTL is known, can be estimated from the total proportion of phenotypic variance explained by all QTL. On the other hand, the dominant nature of AFLP prevented discrimination of allelic composition of the tagged loci in the parental lines, and thereby, analysis of configuration and the effect of different combinations on the trait of interest were not feasible. This is particularly important in polyploid genomes where a greater number of possible alleles at a locus exist.

Despite the drawbacks, the fAFLP data of population Z3 does provide valuable information regarding the genetic components at all tagged loci for each individual of the population Z3. The fAFLP data can therefore be used, at least in population Z3, for informed selection of individuals harbouring superior allele combinations of QTL-linked fAFLP markers based on the origin of the markers and direction of effects on the traits studied.

Studies on function mapping, SNP detection in the functional gene *Stp23*, QTL analysis using candidate genes offer new possibilities for future QTL analysis. On this basis, an improved strategy can be proposed. Since a large number of functional genes have been localized on the potato linkage map, functional genes can be readily selected based on the map positions and used as markers for QTL analysis for any traits of interest. As the genes are predominantly mapped by STS approach, the primer sets can be directly adopted from the present study to generate CAPS markers for the same genes in different populations provided that the same marker alleles are present. When sufficient SNPs are discovered for all these functional genes as done for *Stp23*, the genotyping process could be remarkably accelerated by automated detection and scoring of SNPs.

This strategy could offer many advantages as compared to fAFLP-tagging. First, since STS and SNP markers are potentially codominant in nature and could detect multiple alleles for a given locus, a more powerful program such as interval mapping could be applied to the data for QTL analysis. Second, since the map position of the genes are known, more detailed parameters of QTL, such as the locations, number and degree of effects and the interaction between QTL, could be addressed without constructing a linkage map. Third, when the biochemical basis of a trait under study is highly relevant to carbohydrate metabolism, the candidate gene approach can be used for QTL analysis as performed in the present study.

And finally, once significant QTL are identified, the functional gene markers associated with the trait could be incorporated directly into breeding programs for marker-assisted selection.

### 4.4.2 Magnitude of fAFLP-tagged QTL effects

Based on threshold set for declaring marker-trait associations, many significant QTL were identified in both locations. Some QTL were expressed in one location but not in the other. Some of these potential QTL may be false positives which resulted simply by chance. The frequency of false positives could be tested by permutation tests (Churchill and Doerge 1994) which were, however, not applied in this study. Nevertheless, when a fragment consistently showed association with a trait in independent trials with the same direction of effect or when a fragment showed associations with different but related traits, it was more likely that the fragment is linked to a genuine QTL. Similarly, a QTL detected in single environment may be caused simply by chance. However, considering the stringent threshold *P* value used, most QTL detected in only one location are likely to be the result of genotype x environment interaction.

The amount of variance explained by single fAFLP markers was small. The most significant marker accounted for 6 %, 12.6 % and 6.9 % of the variance of TSC, RSC and FE, respectively. This may be explained by several reasons. First, the genetic complexity of the traits which may be controlled by numerous genes. As described in Sect. 1.6, at least 70 genes are known to be involved in carbohydrate metabolism and transport and may potentially contribute to the biosynthesis of starch and sugars. Many genes may be responsible as QTL for the variation of TSC and RSC depending on the segregation population used (see Sect. 4.2.3). For FE, the knowledge of biochemical mechanisms underlying the trait is not well understood, however the polygenic nature is clearly demonstrated by its phenotypic distribution. There are reports indicating that 17, 6 and 7 QTL for TSC (Schäfer-Pregl *et al.* 1998), chip colour (reflecting the level of RSC) (Douches and Freyre 1994) and foliage maturity (Collins *et al.* 1999), respectively, have been identified in diploid mapping populations. Assuming that a similar number of QTL for TSC with additive effect were active in Z3, given the trait heritability of 59 %, it is conceivable that only 2 % - 6 % of the variation was accounted for by individual markers. Second, the marker density in the tetraploid population may not have been very high, due to the limited number of primer combinations used for marker generation. Consequently, recombination occurred between markers and QTL resulted in the reduction of the effect that can be detected by the marker.

Moreover, the dominant nature of the AFLP markers could not allow more precise classification of genotypes based on their allele configuration except the two classes with marker present or absent. Finally, the variation in trait measurement decreased the heritability of the trait.

### 4.4.3 Assignment of map positions of fAFLPs and segregation analysis

AFLP polymorphisms are defined by nine nucleotides (six nucleotides for the enzyme recognition sites plus three selective nucleotides) at the *EcoRI* end and seven nucleotides (four nucleotides for the enzyme recognition sites plus three selective nucleotides) at the *MseI* end and the fragment size. It is likely that an AFLP fragment with the same size derived from different genetic backgrounds corresponds to the same locus. In contrary to RAPD markers which have an acute problem of transferability (Lynch *et al.* 1994; Black 1993), AFLP fragments mapped in one population can serve as chromosome-specific markers if they also segregate in another population. The transferability of AFLP markers has been demonstrated in barley (Waugh *et al.* 1997; Yin *et al.* 1999) and in potato (Li *et al.* 1998; Rouppe van der Voort *et al.* 1997). This observation provides the basis for inferring map positions of fAFLP markers derived from the tetraploid population Z3 from markers of identical size which are mapped in the reference mapping population K31.

fAFLP was performed on the reference mapping population K31 under the same conditions with the same primer combinations as performed in Z3. 43 fAFLP markers segregating in K31 were detected to be in common between the two populations K31 and Z3. Segregation data created by GeneScan<sup>®</sup> and Genotyper<sup>™</sup> were integrated into the RFLP framework data set of K31 and analysed for linkages. 18 fAFLP markers were localised with a recombination frequency of less than 5 % to the nearest reference markers and multipoint analysis of recombination frequency allowed confident positioning of these markers (Figure 3-10). For the remaining markers, the recombination frequency with the closest reference markers was between 5 % - 18 % which, in most cases, may be overestimated. This may be due to the fact that the adapted threshold of peak quality 50 for fragment calling may not be fully fit for all the cases under investigation. Consequently, this may result in erroneous scores leading to the overestimation of the recombination frequency. Thus these markers were positioned only approximately.



## Discussion

---

The inferred map position of the fAFLP markers, although limited in number, could allow comparison of some fAFLP-tagged QTL with the ones previously detected in other studies. 17 QTL for TSC have been reported in K31 (Schäfer-Pregl *et al.* 1998). Several fAFLP-tagged QTL for TSC in Z3 were positioned in the similar regions. They are G3-393 (LG I), B3-206, G2-71 (LG II), B1-316, G3-342, *SssI* (LG III), G1-90 (LG IV), B2-168 (LG V), B3-121 (LG IX), B3-130 (LG X), *Dbe*, B1-77 (LG XI), B1-103, B3-100, and G2-444 (LG XII). This finding suggests that these markers may have tagged the same gene or gene family controlling TSC as detected in the population K31. The QTL alleles are unlikely, however, the same since Z3 and K31 are genetically unrelated. The genetic variation of the QTL alleles in the two populations may be responsible for the potentially different degree of QTL effects.

Six QTL for chip colour (reflecting the level of RSC) have been revealed (Douches and Freyre 1994) with two on chromosome II, one on IV, two on V and one on X. Some fAFLP markers, were assigned to the same linkage groups such as G1-90 (LG IV), B3-215 (LG V), B3-141 (LG V), G1-342 (LG V), and B3-137 (LG V). Further comparison is not feasible due to the lack of common reference markers between the two populations investigated. Some fAFLP-tagged QTL for RSC were localized in the QTL regions for TSC in K31, such as B1-166 (LG I), G3-83 (LG I), *Stp23* (LG III), G1-90 (LG IV), B2-398 (LG VI), G3-54 (LG IX), *Dbe* (LG XI), G2-89 (LG XI), B1-77 (LG XI), G1-169 (LG XII). The localization of QTL for different traits in a similar regions on molecular linkage maps may be due to pleiotropic effects of single genes on multiple traits. It cannot be excluded, however, that such results may be due to the linkage of two or more functionally independent genes. The lack of a high-resolution map prevents to distinguish these two possibilities.

7 QTL for FE have been identified with the one on LG V linked to marker GP179 being the most significant, accounting for up to 70 % of the phenotypic variation (Collins *et al.* 1999). The functional gene *Sut2*, which was closely linked to GP179, was selected for tagging the QTL in Z3 (see Sect.3.3.6). No effect was detected at this locus for FE in Z3. It should be noted that 5 fAFLP markers showing associations with FE in Z3 were clustered in an 18 cM interval flanked by RFLP markers GP179 and GP234. These multiple markers proximal to the region reinforce that the markers tagged a genuine QTL for FE. However, no fAFLP markers in this region showed such a notably significant effects as detected in the population analysed by Collins *et al.* (1999). One explanation is that QTL alleles in this population are different from that in Z3. It is also possible that the fAFLPs tagged a closely linked but

## Discussion

---

functionally different gene(s), or breeders' selection against very late genotype diminished the effects.

The inferred map position of the fAFLP markers represent only a small part of the markers showing associations with the analysed traits. More detailed QTL parameters in terms of location, number, interaction (epistasis), reproducibility and degree of effect, would rely on the availability of map positions of all other markers. This could be compensated using various reference mapping populations. An initial step has been made by Rouppe van der Voort *et al.* (1998) for AFLP mapping in potato by which a number of AFLP markers have been positioned on different linkage groups. When more and more AFLP markers with their map positions become available, it may become more feasible to infer the position of a AFLP marker from a reference AFLP map without cloning and sequencing.

The possible segregation type of fAFLP markers was determined based on a chi-square test for deviation from the segregation ratios of 1:1 and 1:5 expected for simplex- nulliplex and duplex-nulliplex fragment configurations, respectively. Segregation ratios of 1:3, 1:11 and 1:35 are expected for parental genotype of simplex-simplex, simplex-duplex and duplex-duplex, respectively. 49 markers showed the expected segregation ratio as indicated in Tables 3-8, 3-9, 3-10. However, caution must be taken when reviewing the results since the method of inferring the mode of segregation type may result in a false hypothesis when segregation ratios are highly distorted. In most cases, distorted segregation ratios that were observed mainly favoured the undetected alleles (absence). Some may be due to the occurrence of double reduction which increases the production of homozygous gametes as compared to random chromosome segregation and leads to alteration of segregation ratios (Ronfort *et al.* 1998), others may be explained by the selection performed on the population by breeders. Distortion favouring the presence of a marker allele may also exist but cannot be tested reliably in the study because of the multiple alternative genetic models and the limited number of individuals available for testing, for example, goodness of fit for the ratio 1:35.

### 4.4.4 Genetic basis of correlations between phenotypic traits

A negative correlation ( $r=-0.62$ ) was found for TSC and RSC in the population Z3. This relationship was also uncovered from QTL analysis. 15 marker loci were detected as having associations with both traits (Table 3-10). At each of these fAFLP loci, except G1-339, the alleles associated with increased TSC were associated with decreased RSC and *vice versa*.

## Discussion

---

This close association between QTL for TSC and RSC at these marker loci suggests that both traits may be controlled at least in part by the same gene(s) with pleiotropic effects.

Similarly, TSC and FE were also correlated with a correlation coefficient of 0.49 (see Sect. 3.3.2). Generally, high TSC is associated with delayed foliage maturity and *vice versa*. This was further confirmed in the QTL analysis by which 12 markers showed associations with TSC and FE. All markers except Y-393 and B1-244 had effects on both traits with direction congruent to their phenotypic correlation relationship, indicating that the same gene(s) with pleiotropic effects may be responsible for the QTL identified.

Fewer markers (7) have been detected to be associated with both RSC and FE, and the direction of effect on the two traits, unlike the previous two cases, was variable (Table 3-10). 3 markers were accordant in the direction of effect with the phenotypic correlation while 4 markers were contradictory. The variability of the direction may indicate the genetic complexity of the two traits. Some fAFLP may have tagged genetic factors with pleiotropic effects while others may detect functionally different but closely linked genes.

Overall, the marker-trait association analysis has revealed a number of fAFLP markers associated with single and multiple traits. The relationship between traits revealed by markers with multiple effects on the traits is consistent with their phenotypic correlations indicating that the fAFLP markers successfully tagged, at least in part, genetic factors contributing to the traits examined. The QTL analysis with fAFLP markers provide a genetic basis in terms of linkage or pleiotropy for the correlations observed between these traits. Briefly, the traits relationship between TSC and RSC, TSC and FE may be in part controlled by the same genetic factors with pleiotropic effects on different traits. It cannot be excluded, however, that functionally different but genetically closely linked genes are the cause of the correlations. Further work is required to get the definitive conclusions on this point.

The nature of pleiotropic effects of QTL on TSC and RSC was in agreement with the expectation of potato breeding for high TSC and low RSC. The partially common or closely linked genetic factors for TSC and FE imply that it may be difficult to increase TSC without delaying the foliage maturity through conventional breeding procedures. In this case, the use of molecular markers for the simultaneous improvement of TSC and FE may be more efficient than direct selection based on the phenotype. This could be achieved by pyramiding favourable alleles for both traits at independent QTL.

### 4.4.5 Cloning and conversion of fAFLP markers and comparison of the two genotyping systems fAFLP and rAFLP

The marker-trait association analyses yield QTL being tagged by a number of fAFLP markers. The most significantly associated fAFLP markers need to be cloned and converted into easy-to-use markers for further investigation. To this end, a case study of cloning fAFLP fragments via radioactive AFLP was carried out using primer combination of B3. The experiment has demonstrated that it was possible to clone fAFLP markers via rAFLP approach. However, the efficiency of cloning was relatively low. There are reports indicated that successful cloning and conversion of AFLP markers was mostly limited to a range of smaller AFLP fragments (Reamon-Buettner and Jung 2000). In present study, cloning of fAFLP was bypassed using rAFLP. This technique change resulted in further decreasing of cloning efficiency since not all fragments could be re-identified unambiguously on autoradiograph. The conversion of the marker proved to be much more difficult than expected which could be attributed to the tetrasomic genome structure and the nature of the AFLP fragments. First, sequencing of ten clones revealed 7 nearly identical sequences which differed from one another by single nucleotides. The sequences were highly variable and the actual number of such sequences is likely more. Second, the polymorphism of the fAFLP fragment is based on the *Eco*RI and *Mse*I sites and the three selective nucleotides of the primers. Primers that are designed in any other region may result in lost of specificity defined by AFLP (Shan *et al.* 1999). Third, allele specificity revealed by sequence comparison between two parents could be useful for designing allele-specific primers for diploid inbred species but may not be so in highly heterozygous tetraploids like potato. This is further complicated by the fact that the cloned fragment detected multiple copies. Taking all these factors into consideration, conversion of fAFLP markers is a technically difficult and time-consuming exercise.

Cloning of fAFLP fragments via rAFLP also allowed the comparison of the two genotyping systems. Compared to rAFLP, fAFLP offers many advantages. (1) fAFLP is performed with fluorescence-labelled primers in combination with DNA sequencer, thus precluding the necessity of working with radioactivity. (2) Since primers are labelled with different fluorescent dyes, fAFLP products could be mixed and loaded in one lane on the gel, thus significantly reducing the number of gels to be run per study. (3) Electrophoresis of fAFLP is viewable directly on the gel image at real time, and the process can be completed within 4 hours using appropriate running parameters. The results that can be subjected to down stream

## Discussion

---

computer-aided analysis eliminating all steps like fixing, drying and exposure as required by rAFLP analysis. (4) Highly precise sizing with single base pair resolution and data output both qualitative (scored as presence and absence) and quantitative (scored as peak height) in *Excel* format can be achieved semi-automatically by GeneScan® and Genotyper™ software eliminating the possibility of errors occurring during manual data handling. (5) fAFLP generates fragments with a larger scorable size range from 50bp to 500bp or more if needed, while fragments outside the range of 100-500 bp were hardly scorable on autoradiographs. (6) fAFLP is extremely sensitive and powerful in detecting trace amount of product while for rAFLP, the trace products appear as faint bands which vary with the exposure time and the decay period of radioactivity of the label used. Therefore, fAFLP generated much more scorable fragments than rAFLP. For example, a total of 117 scorable fragments were generated with primer pair E-act/M-cta using the automated fluorescent AFLP system, whereas only 72 fragments could be clearly identified on autoradiograph using radioactively labelled AFLP technique. This discrepancy is largely due to the faint bands that could hardly be unambiguously scored on autoradiographs. Similar observations were also reported by Zhang *et al.* (1999).

While having these advantages, fAFLP technology has also inherent problems. Since it is extremely sensitive and detects even trace amounts of product, it can also easily generate environmental noise resulting from either the overlapping of the wavelengths between different colour dyes or the contamination of gel plates. To minimize the influence on the results, a quality threshold of 50 was selected for fragment scoring. In other words, only fragments that were higher than 50 in peak height were scored. In my experience, such threshold seemed optimal although it was chosen arbitrarily. Moreover, the overlapping of the wavelength between different dyes can cause spectral interference between the dye labels during analyses. Consequently, the size standard definition derived from one of the samples may not fully fit to those of others, which, when not verified, may result in an erroneous assignment of size to fAFLP fragments. Therefore, the size standards of all the samples have to be visually checked to eliminate this possibility before proceeding to data analysis with Genotyper™. In genotyping using Genotyper™ software, fragment calling is performed by converting fragment size into prior-defined size categories. In the present study, the range of size categories was limited to a single base pair. As a result, the same-size fragments may be split into two different ones in some cases due to slight lane-to-lane variation. A post viewing of the output data is also required to minimize such errors. Finally, recovering of fAFLP

fragments of interest cannot be performed on the system itself to date but has to rely to rAFLP. Due to the difference of sensitivity of the two genotyping systems, rAFLP could not detect all the fragments which are identified on the fAFLP system.

### 4.4.6 Function map-based QTL analysis

The candidate gene strategy has been shown to be very useful and effective in the analysis of genetic factors controlling phenotypic traits with a known biochemical and physiological basis (Winkelman 1992, Causse *et al.* 1995, Faris *et al.* 1999). The basis of the candidate gene strategy is that relationships between functional genes and traits are known and that genome regions showing a QTL effect in one population also show marker-QTL associations in the same region in other populations (Oberhagemann *et al.* 1999, Kearsey and Farquhar 1998). In potato, the function map for genes involved in carbohydrate metabolism (see Sect. 3.1) provides the framework for the application of candidate gene strategy in the QTL analysis for starch and RSC. Five functional genes selected on the basis of known overlapping positions with QTL for TSC identified in population K31 (Schäfer-Pregl *et al.* 1998) were tested for association with phenotypic traits in population Z3. Association analysis indicated that genes *Dbe* and *SssI* were significantly associated with QTL affecting TSC in population Z3 at the 1 % level. No significant effect was detected at the other gene loci. More prominent results were obtained for RSC. Two loci *Stp23* and *Sut2*, on chromosomes III and V, respectively, showed significant effects consistently detected at two locations at 1 % level. It is worth noting that *Stp23* is closely linked to *Pain-1* (coding for soluble acid invertase, see Figure 3-3 on page 57) which, based on its functional role, appears more likely to be the candidate for the QTL controlling RSC. The effects observed at *Stp23* may result from the linked *Pain-1* gene. Gene *Sut2*, a member of sucrose transporter family, was repeatedly detected as a QTL affecting RSC. The observation is in accordance with the putative physiological role of *Sut2*, which is believed to act as sucrose sensor to control sucrose fluxes across the plasma membrane of sieve element (Barker *et al.* 2000). It should be noted that a major QTL for FE was identified at RFLP locus GP179 (Oberhagemann *et al.* 1999, Collins *et al.* 1999), however, no such effects were detected at the *Sut2* locus, which is closely linked to GP179, in the population Z3. The possible explanations may be that the effects were not large enough to be detected in the experiment with the statistical method used or alleles were homozygous at putative QTL loci or a PCR-based assay at locus *Sut2* may have detected a allele ineffective on the trait.

## 5. SUMMARY

---

### 5.1 Function mapping as an approach to the identification of genes controlling tuber starch content in potato

Carbohydrates are the major storage compounds of cereal and tuber crops. In potato, tuber starch content, which is a quantitative character, is of agronomic importance and is controlled by genetic factors and environmental conditions. Genetic mapping using anonymous RFLP markers allowed the dissection of the genetic component of the quantitative trait into discrete quantitative trait loci (QTL) and their localization on the potato molecular map. QTL analysis alone does not reveal the identity of the gene(s) controlling the trait at the molecular level. Molecular linkage maps based on functional gene markers may shed light on this issue by correlating the relevant functional genes to QTL for the trait, thus leading to the identification of candidate genes underlying a quantitative trait. For this purpose, the nucleotide sequence databases EMBL and GenBank were explored for sequence information of genes from potato, tomato and other plant species. Priority was given to genes operating in carbohydrate metabolism and transport. Sequence information was used to amplify the homologous or analogous genes in potato by PCR. Polymorphisms of the amplification products were revealed either directly or after cleavage with restriction enzymes in two diploid mapping populations K31 and LH. Reliability of the mapping approach was confirmed by sequence comparison of the PCR products with the source genes and by comparing map positions of PCR products with corresponding loci previously determined by RFLP. A potato function map consisting of 82 loci was constructed based on 66 cloned genes of known biochemical function. The comparison of the molecular function map with the QTL map for tuber starch content, which was previously constructed using the same population, revealed a number of correlations between the map positions of QTL for tuber starch content and loci of function-related genes. 14 out of 17 QTL for tuber starch content were linked to at least one known candidate gene locus. The clearest positional correlation was found between the *Stp23* gene encoding starch phosphorylase and QTL *ts(g)* on linkage group III.

The gene markers, as defined in this study, are a valuable tool for genetic analysis and marker-assisted selection since they are predominantly PCR-based and cover a considerable

proportion of the potato genome. To enhance further the efficiency of genotyping a population for a given functional gene, it is highly desirable that the gene markers are amenable to automatization. To this end, candidate gene *Stp23* was selected as a case study for exploring single nucleotide polymorphisms (SNPs), which have much potential for automated detection and scoring. PCR products of 14 potato lines were sequenced. Bi-allelic and tri-allelic SNPs were revealed within single lines and sequence assembly of different lines revealed potential SNPs existing in potato genotypes. Together, 31 candidate SNPs were identified in a 1801 bp segment of *Stp23*, 8 of which were in coding regions and 23 were in introns.

## **5.2 fAFLP tagging of genetic factors controlling tuber starch content, reducing sugar content and foliage earliness in tetraploid potato**

Tuber starch content, reducing sugar content and foliage earliness are three important agronomic traits in potato breeding program. QTL analysis for the traits has been addressed in a few studies using diploid potato. While providing insight into the genetic nature of the trait studied, studies performed on diploids may suffer from the disadvantage of ploidy manipulation and difference in QTL effects may exist at two ploidy levels (Groose *et al.* 1988). Therefore in the second section of the study, a QTL analysis was carried out in collaboration with a breeding company in a tetraploid hybrid family Z3. An efficient fAFLP genotyping system performed on a 377 DNA sequencer was adopted for DNA marker generation. A total of 806 fAFLP markers were tested for association with QTL for tuber starch content, reducing sugar content and foliage earliness. 40, 25, 32 fAFLP fragments were identified to be associated with tuber starch content, reducing sugar content and foliage earliness, respectively. 33 fAFLPs showed associations with two or more traits.

To obtain positional information of these fAFLP markers on the potato map, the reference mapping population K31 was also genotyped in parallel with the same system and using the same primer combinations. Map positions of 50 fAFLP fragments were inferred from common markers mapped in the reference mapping population K31. Nine were located in map regions of previously identified QTL for tuber starch content.



## Summary

---

For breeding-oriented QTL analysis, the most significantly associated markers are required to be cloned and converted into a marker type that is amenable to marker-assisted selection. To achieve this goal, radiolabelled AFLP was used for cloning fAFLP fragments of interest. However, cloning and conversion of fAFLP fragments via rAFLP was not efficient due to the difference in sensitivity between the two genotyping systems and the complexity of tetrasomic inheritance of tetraploid potato. Only one fragment, which was highly associated with reducing sugar content, was cloned and sequenced but failed to be converted into a 'easy-to-use' marker.

With the availability of the potato function map described in the first section, it became feasible to perform QTL analysis via a candidate gene approach for any given population. To test this, five candidate genes, which overlapped with QTL for tuber starch content on the map of K31, were selected to test for association with the traits investigated in population Z3. Genes *Stp23* (starch phosphorylase) and *Sut2* (sucrose transporter like) detected significant associations with reducing sugar content at two different locations. Three alleles of the gene *Dbe*, coding for a debranching enzyme, had an effect on tuber starch content at one location.

## 5. ZUSAMMENFASSUNG

---

### 5.1 Funktionelle Kartierung als ein Ansatz zur Identifikation von Genen, die den Stärkegehalt der Kartoffelknolle kontrollieren

Kohlenhydrate sind in Getreidepflanzen und knollenbildenden Feldfrüchten die dominierenden Speicherverbindungen. Der Stärkegehalt der Kartoffelknolle ist ein quantitatives Merkmal von agronomischer Bedeutung und wird durch genetische Faktoren und Umwelteinflüsse reguliert. Dieses quantitative Merkmal konnte durch genetische Kartierung mit Hilfe von anonymen RFLP Markern in diskrete QTLs zerlegt werden, die auf der genetischen Karte der Kartoffel lokalisiert wurden. Die QTL-Analyse allein führt hingegen nicht zur Identifizierung der merkmalskontrollierenden Gene auf der molekularen Ebene. Hier können molekulare Kopplungskarten hilfreich sein, die auf funktionellen Gen-Markern basieren. Die Korrelation von Merkmals-relevanten Genen mit QTLs für dieses Merkmal kann zur Identifikation von Kandidatengenen führen, die das quantitative Merkmal prägen. Zu diesem Zweck wurden die Sequenzdatenbanken EMBL und GenBank vorwiegend nach Kartoffel- und Tomatengenen, aber auch Genen anderer Pflanzen durchsucht, die dem Kohlenhydratstoffwechsel und -transport zuzuordnen sind. Von der Sequenzinformation dieser Gene wurden Primer abgeleitet, mit deren Hilfe homologe bzw. analoge Sequenzen aus der Kartoffel mittels PCR amplifiziert wurden. Segregierende Polymorphismen zeigten sich nach der Amplifikation von DNA der verschiedenen Pflanzen der diploiden Kartierungspopulationen K31 und LH, entweder direkt oder nach dem Restriktionsverdau der PCR-Produkte. Die Verlässlichkeit dieses Kartierungsansatzes wurde durch Vergleich der Nukleotidsequenz der PCR-Produkte mit der Datenbanksequenz, von der die Primer abgeleitet worden waren, und durch Vergleich der Positionen der gewonnenen PCR-Produkte auf der Genkarte mit den entsprechenden bereits vorher mittels RFLP bestimmten Loci überprüft. Hierdurch wurde eine funktionelle Genkarte der Kartoffel mit 82 Loci basierend auf 66 klonierten Genen von bekannter biochemischer Funktion erstellt. Die QTL-Karte des Kartoffelgenoms bezüglich des Stärkegehalts der Kartoffelknolle, welche von derselben Population konstruiert worden war, konnte nun mit der funktionelle Genkarte verglichen werden. Es wurde deutlich, daß sich eine bedeutende Anzahl an Positionen aus den beiden verschiedenen Kartierungsansätzen überlagerten. 14 der 17 QTLs konnten mindestens einem Locus für ein bekanntes Kandidaten-Gen zugeordnet werden. Die beste positionelle

Korrelation fand sich zwischen dem *Stp23*-Gen, das für eine Stärkephosphorylase codiert, und dem QTL *ts(g)* auf Chromosom III.

Die von bekannten Genen abgeleiteten Marker, wie sie in dieser Arbeit vorgestellt werden, sind ein nützliches Hilfsmittel für die genetische Analyse und für die Marker gestützte Selektion, da sie hauptsächlich auf PCR-Techniken basieren und einen beträchtlichen Teil des Kartoffelgenoms abdecken. Zur Steigerung der Effizienz der genotypischen Analyse einer Population bezüglich eines gegebenen funktionellen Gens ist es wünschenswert, den Gen-Marker einer automatischen Analyse zugänglich zu machen. SNPs (single nucleotide polymorphisms) eignen sich besonders gut für eine automatische Detektion und Auswertung. Beispielhaft wurden hierzu PCR-Produkte des Kandidaten-Gens *Stp23* von 14 Kartoffellinien sequenziert. Innerhalb einzelner Kartoffellinien wurden bi- und triallelische SNPs gefunden. Der Vergleich der Sequenzen verschiedener Genotypen führte zur Identifizierung potentieller, für die Automatisierung geeigneter SNPs. So wurden insgesamt 31 SNPs in einem 1801 bp langen Segment des *Stp23*-Gens identifiziert von denen 8 in der codierenden Region und 23 in den Introns zu finden waren.

## **5.2 fAFLP-Tagging von genetischen Faktoren, die den Stärkegehalt, den Gehalt von reduzierenden Zuckern der Kartoffelknolle und die Reifezeit kontrollieren**

In der Kartoffelzüchtung sind der Stärkegehalt der Knollen, der Gehalt an reduzierenden Zuckern bei der Lagerung und die Reifezeit Merkmale von herausragender Bedeutung. Einige Studien widmen sich der QTL-Analyse dieser agronomischen Merkmale unter Verwendung von diploiden Kartoffellinien. Diese liefern zwar einen aufschlußreichen Einblick in die genetische Natur der untersuchten Merkmale, reichen jedoch nicht zur Beschreibung von QTL-Effekten bei tetraploiden Kartoffellinien. Deshalb wurde im zweiten Teil dieser Arbeit in Zusammenarbeit mit einem Kartoffelzuchtunternehmen eine QTL-Analyse einer tetraploiden Hybridfamilie Z3 unternommen. Zur Gewinnung von DNA-Markern wurden auf effizient Weise fAFLP Produkte mit einem 377 DNA-Sequenzierer analysiert. Insgesamt wurden 806 fAFLP Marker auf ihre Kopplung mit QTLs für den Stärkegehalt der Kartoffelknolle, den Gehalt an reduzierenden Zuckern und die Reifezeit hin untersucht. Davon waren 40 fAFLP Marker mit QTL für den Stärkegehalt, 25 mit QTL für

den Gehalt an reduzierenden Zuckern und 32 mit QTL für die Reifezeit assoziiert. 33 fAFLP Marker waren mit zwei oder mehr Merkmalen gekoppelt.

Um positionelle Information über diese fAFLP Marker auf der genetische Karte der Kartoffel zu erhalten, wurde auch die diploide Kartierungspopulation K31 mit derselben Technik unter Verwendung der gleichen Primerkombinationen untersucht. 50 fAFLP Fragmente waren mit bereits bekannten RFLP-Markern in der Referenzpopulation K31 gekoppelt. 9 fAFLP-Fragmente wurden in Regionen lokalisiert, die zuvor als QTLs für den Stärkegehalt der Kartoffelknolle identifiziert worden waren.

Will man die signifikantesten Marker für eine QTL-Analyse in einem Züchtungsprogramm nutzen, so müssen diese kloniert und in einen Markertyp umgewandelt werden, der in der Marker-gestützten Selektion leicht zu handhaben ist. Hierzu wurden fAFLP-Fragmente in radioaktiv markierte rAFLP-Fragmente überführt. Es stellte sich allerdings heraus, daß die Klonierung und Umwandlung von fAFLP-Fragmenten über die Herstellung von rAFLP-Fragmenten nicht effizient genug ist. Einerseits ist die Sensitivität der beiden Markierungstechniken der AFLP-Fragmente zu unterschiedlich, andererseits begrenzt die Komplexität der Vererbung der genetischen Faktoren bei tetraploiden Kartoffeln eine problemlose Umwandlung aller fAFLP-Fragmente in radioaktiv markierte, klonierbare AFLP-Produkte. Allerdings gelang es, ein AFLP-Fragment, das hoch signifikant mit dem Gehalt an reduzierenden Zuckern gekoppelt war, zu klonieren und zu sequenzieren. Ein einfach zu handhabender Marker konnte von diesem Fragment jedoch noch nicht abgeleitet werden.

Mit der Verfügbarkeit einer funktionellen Karte des Kartoffelgenoms, so wie sie im ersten Abschnitt beschrieben wurde, ist es möglich, eine QTL Analyse mittels Nutzung der Sequenzinformation von Kandidaten-Genen für jede beliebige Population durchzuführen. Zum Test wurden fünf Kandidaten-gene, die mit QTLs für den Stärkegehalt der Kartoffelknolle in der Population K31 gekoppelt sind, auch in der Population Z3 auf Assoziation mit dem Stärkegehalt hin untersucht. Drei Allele des *Dbe*-Gens, das für das "debranching"-Enzym kodiert, zeigten einen Effekt auf den Stärkegehalt. Die Gene *Stp23* (Stärkephosphorylase) und *Sut2* (Saccharosetransporter-ähnlich) zeigten eine signifikante Kopplung mit dem Gehalt an reduzierenden Zuckern.

## 6. REFERENCES

---

- Abel GJ, Springer F, Willmitzer L and Kossmann J** (1996) Cloning and functional analysis of a cDNA encoding a novel 139 kDa starch synthase from potato (*Solanum tuberosum* L.) Plant J. 10 (6):981-991
- Ajmone-Marsan P** (1997) AFLP markers for DNA fingerprinting in cattle. Anim. Genet. 28:418-426
- Akkaya MS, shoemaker RC, Specht JE, Bhagwat AA, Cregan PB** (1995) Integration of simple sequence repeat (SSR) DNA markers into a soybean linkage map. Crop Sci. 35:1439-1445
- Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, Van den Bosh N, Pot J, and Kuiper MTR** (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. Plant J. 14:259-271
- Austin DF, Lee M** (1996) comparative mapping in F2:3 and F6:7 generations of quantitative trait loci for grain yield and yield components in maize. Theor. Appl. Genet. 92:817-826
- Bachem CWB, Vanderhoeven RS, Debruijn SM, Vreugdenhil D, Zabeau M, Visser RGF** (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP-analysis of gene expression during potato tuber development. Plant J. 9(5):745-753
- Barone A, Ritter E, Schachtshabel U, Debener T, Salamini F, Gebhardt C** (1990) Localisation by restriction fragment length polymorphism mapping in potato of a major dominant gene conferring resistance to the potato cyst nematode *Globodera rostochiensis*. Mol. Gen. Genet. 244:177-182
- Barker L, Kühn C, Weise A, Schulz A, Gebhardt C, Hirner B, Hellmann H, Schulze W, Ward JM and Frommer WB** (2000) A sucrose transporter homolog in sieve elements is a candidate gene for QTLs controlling starch and yield. (in press)
- Beckmann JS and Soller M** (1983) Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. Theor. Appl. Genet. 67:35-43
- Becker J, Vos P, Kuiper M, Salamini F, Heun M** (1995) Combined mapping of AFLP and RFLP markers in barley. Mol. Gen. Genet. 249:65-73
- Bell CJ and Ecker JR** (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19:137-144
- Bendahmane A, Kanyuka K, Baulcombe D** (1997) High resolution genetic and physical mapping of the Rx gene for extreme resistance to potato virus X in tetraploid potato. Theor. Appl. Genet. 95:153-162
- Bernacchi D, Beckbunn T, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley S** (1998) Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *L. hirsutum*. Theor. Appl. Genet. 97(3):381-397
- Bernacchi D, Beckbunn T, Eshed Y, Inai S, Lopez J, Petiard V, Sayama H, Uhlig J, Zamir D, Tanksley S** (1998) Advanced backcross QTL analysis in tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *L. hirsutum* and *L. pimpinellifolium*. Theor. Appl. Genet. 97(1-2):170-180
- Bernatzky R, Tanksley SD** (1986a) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887-898
- Bevan M, Bancroft I, Bent E, Love K, Goodman H, Dean C, Bergkamp R, Dirkse W, Van Staveren M, Stiekema W, Drost L, Ridley P, Hudson S-A, Patel K, Murphy G, Piffanelli P, Wedler H, Wedler E, Wambutt R, Weitzenegger T, Pohl TM, Terryn N, Gielen J, Villarroel R, De Clerck R, Van Montagu M, Lecharny A, Auborg S, Gy I, Kreis M, Lao N, Kavanagh T, Hempel S, Kotter P, Entian K-D, Rieger M, Schaeffer M, Funk B, Mueller-Auer S, Silvey M, James R, Montford A, Pons A,**

## References

---

- Puigdomenech P, Douka A, Voukelatou E, Mlioni D, Hatzopolous P, Piravandi E, Obermaier B, Hilbert H, Düsterhöft A, Moores T, Jones JDG, Eneva T, Palme K, Benes V, Rechman S, Ansorge W, Cooke R, Berger C, Delseny M, Voet M, Volckaert G, Mewes HW, Klosterman S, Schueller C, Chalwatzis N** (1998) Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* 391: 485-488
- Black WC IV** (1993) PCR with arbitrary primers: approach with care. *Insect Mol. Biol.* 2:1-6
- Blakeley SD, Plaxton WC and Dennis DT** (1990) Cloning and characterization of a cDNA for the cytosolic isozyme of plant pyruvate kinase: the relationship between the plant and non-plant enzyme. *Plant Mol. Biol.* 15 (4):665-669
- Boivin K** (1999) Towards a saturated sorghum map using RFLP and AFLP markers. *Theor. Appl. Genet.* 98:320-328
- Bonierbale M, Plaisted RL, Tanksley SD** (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics* 120:1095-1103
- Bonierbale MW, Plaisted RL, Pineda O, and Tanksley SD** (1994) QTL analysis of trichome-mediated resistance in potato. *Theor. Appl. Genet.* 87:973-987
- Borchert S, Harborth J, Schunmann D, Hoferichter P, and Heldt HW** (1993). Studies of the enzymic capacities and transport properties of pea root plastids. *Plant Physiol.* 101:303-12
- Botstein D, White RL, Skonick M, Davis RW** (1980) Construction of a genetic linkage map using restriction fragment length polymorphism. *Amer. J. Human. Genet.* 32:314-331
- Bradshaw JE, Mackay GR** (1994) Breeding strategies for clonally propagated potatoes. In: Bradshaw JE, Mackay GR (eds) *Potato genetics*. CAB international Wallingford, UK, pp 467-497
- Bradshaw JE, Meyer RC, Milbourne D, McNicol JW, Phillips MS, Waugh R** (1998) Identification of AFLP and SSR markers associated with quantitative resistance to *Globodera pallida* (stone) in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*) with a view to marker-assisted selection. *Theor. Appl. Genet.* 97(1-2):202-210
- Brigneti G, Garcia-Mas J, Baulcombe D** (1997) Molecular mapping of the potato virus Y resistance gene *Ry-sto* in potato. *Theor. Appl. Genet.* 94:198-203
- Brookes AJ** (1999) The essence of SNPs. *Gene* 234:177-186
- Brunel D, Froger N, and Pelletier G** (1999) Development of amplified consensus genetic markers (ACGM) in *Brassica napus* from *Arabidopsis thaliana* sequences of known biological function. *Genome* 42:387-402
- Burr B, and Burr FA** (1991) Recombinant inbreds for molecular mapping in maize. *Trends Genet.* 7:55-60
- Byrne PF, McMullen MD, Snook ME, Musket TA, Theuri JM, Widstrom NW, Wiseman BR, Coe EH** (1996) Quantitative trait loci and metabolic pathways: genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks. *Pro. Natl. Acad. Sci. USA* 93:8820-8825
- Byrne PF, McMullen MD, Wiseman BR, Snook ME, Musket TA, Theuri JM, Widstrom NW, Coe EH** (1998) Cell biology and Molecular genetics -- maize silk maysin concentration and corn earworm antibiosis: QTL and genetic mechanisms. *Crop Sci.* 38:462-471
- Calvin M and Bassham JA** (1962) *The Photosynthesis of Carbon compounds*, Benjamin, New York
- Carlisle SM, Blakeley SD, Hemmingsen SM, Trevanion SJ, Hiyoshi T, Kruger NJ and Dennis DT** (1990) Pyrophosphate-dependent phosphofructokinase. Conservation of protein sequence between the alpha- and beta-subunits and with the ATP-dependent phosphofructokinase. *J. Biol. Chem.* 265 (30):18366-18371
- Causse M, Rocher JP, Henry AM, Charcosset A, Prioul JL, and Vienne D** (1995) Genetic dissection of the relationship between carbon metabolism and early growth in maize, with emphasis on key-enzyme loci. *Mol. Breed.* 1:259-272

## References

---

- Chang C, Bowman JL, De John AW, Lander ES, Meyerowitz EW** (1988) Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Pro. Natl. Acad. Sci. USA* 85:6856-6860
- Chao S, Baysdorfer C, Heredia-Diaz O, Musket T, Xu G, Coe EH** (1994) RFLP mapping of partially sequenced leaf cDNA clones in maize. *Theor. Appl. Genet.* 88: 717-721
- Cho RJ, Mindrinos M, Richards DR, Sapolsky RJ, Anderson M, Drenkard E, Dewdney L, Reuber TL, Stammers M, Federspiel N, Theologis A, Yang WH, Hubbell E, Au M, Chung EY, Lashkari D, Lemieux B, Dean C, Lipshutz RJ, Ausubel FM, Davis RW, Oefner PJ** (1999) Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nature Genetics.* 23(2):203-207
- Cho YG, Blair MW, Panaud O, McCouch SR** (1996) Cloning and mapping of variety-specific rice genomic DNA sequences-amplified fragment length polymorphisms (AFLP) from silver-stained polyacrylamide gels. *Genome* 39:373-378
- Cho YG, McCouch SR, Kuiper M, Kang MR, Pot J, Groenen JTM, Eun MY** (1998) Integrated map of AFLP, SSLP and RFLP markers using a recombinant inbred population of rice (*Oryza sativa* L) *Theor. Appl. Genet.* 97(3):370-380
- Churchill GA, Doerge RW** (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138:963-971
- Collins A, Milbourne D, Ramsay L, Meyer R, Chatot-Balandras C, Oberhagemann P, De Jong W, Gebhardt C, Bonnel E, Waugh R** (1999) QTL for field resistance to late blight in potato are strongly correlated with maturity and vigour. *Molecular Breeding* 5(5):387-398
- Collins FS, Brooks LD, Chakravarti A** (1998) A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res.* 8:1229-1231
- Corte-Real HB, Dixon DR, and Holland PW** (1994) Intron-targeted PCR: a new approach to survey neutral DNA polymorphism in bivalve populations. *Mar. Biol.* 120:407-413
- Crane MB, Lawrence WJC** (1929) Genetical and cytological aspects of incompatibility and sterility in cultivated fruits. *J. Pomol. Hortic. Sci.* 7: 276-301
- Cregan PB, Bhagwat AA, Akkaya MS, Jiang Rongwen** (1994) Microsatellite fingerprinting and mapping of soybean. *Methods Mol. Cell Biol.* 5:49-61
- Cregan PB, Mudge J, Fickus EW, Danesh D, Denny R, Young ND** (1999) Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by the *rhg1* locus. *Theor. Appl. Genet.* 99(5):811-818
- De Jong W, Forsyth A, Leister D, Gebhardt C, Baucomb D** (1997) A potato hypersensitive resistance gene against potato virus X maps to a resistance gene cluster on chromosome 5. *Theor. Appl. Genet.* 95:246-252
- De Nettancourt D** (1977) Incompatibility in angiosperms. In *Monographs on Theoretical and Applied Genetics*, No. 3. R. Frankel, G.A.E. Gall, and H.F. Linkkens, eds. (Berlin: Springer-Verlag).
- Delseny M, Cooke R, Raynal M, and Grellet F** (1997) The *Arabidopsis thaliana* cDNA sequencing projects. *FEBS Lett.* 403:221-224.
- Dennis DT, Huang Y, Negm FB** (1997) Glycoysis, the pentose phosphate pathway and anaerobic respiration. In: *Plant Metabolism* pp105-115, 2<sup>nd</sup> editon, ISBN 0-582-25906-1, Addison Wesley Longman Limited 1997
- Devos KM, Gale MD** (1992) The use of random amplified polymorphic DNA markers in wheat. *Theor. Appl. Genet.* 84:567-572
- Donini P, Elias ML, Bougourd SM, and Koebner RMD** (1997) AFLP fingerprinting reveals pattern differences between template DNA extracted from different plant organs. *Genome* 40:521-526
- Douches DS, Freyre R** (1994) Identification of genetic factors influencing chip colour in diploid potato (*Solanum tuberosum*). *Am. Potato J.* 71:581-590

## References

---

- Du Jardin P, Rojas-Beltran J, Gebhardt C, Brasseur R** (1995) Molecular cloning and characterization of a soluble inorganic pyrophosphatase in potato. *Plant Physiol.* 109: 853-860
- Eckes P, Schell J, Willmitzer L** (1985) Organ-specific expression of three leaf/stem specific cDNAs from potato is regulated by light and correlated with chloroplast development. *Mol. Gen. Genet.* 199: 216-224
- Edwards A, Marshall J, Sidebottom C, Visser RG, Smith AM and Martin C** (1995) Biochemical and molecular characterization of a novel starch synthase from potato tubers. *Plant J.* 8 (2):283-294
- El-Kharbotly A, Leonard-Schippers C, Huigen D, Jacobsen E, Pereira A, Stiekema W, Salamini F, Gebhardt C** (1994) Segregation analysis and RFLP mapping of the R1 and R3 alleles conferring race specific resistance to *Phytophthora infestans* in progeny of dihaploid potato parents. *Mol. Gen. Genet.* 242:749-754
- El-Kharbotly A, Palomino-Sanchez C, Salamini F, Jacobsen E, Gebhardt C** (1996a) Race-specific resistance against *Phytophthora infestans* in potato is controlled by more genetic factors than only R-genes. *Euphytica* 90:331-336
- Falconer DS** (1960) *Introduction to Quantitative Genetics*. 1<sup>st</sup> Edn, Oliver and Boyd, Edinburgh
- Faris JD, Li WL, Liu DJ, Chen PD, Gill BS** (1999) Candidate gene analysis of quantitative disease resistance in wheat. *Theor. Appl. Genet.* 98:219-225
- Feinberg AP, Vogelstein B** (1983) A technique of radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13
- Feinberg AP, Vogelstein B** (1984) A technique of radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 137:266-267
- Fisher RA** (1918) The correlations between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Ediub.* 52:399-433
- Flügge UI, and Heldt HW** (1991) Metabolite translocator of the chloroplast envelope. *Ann. Rev. Plant Physiol. Mol. Biol.* 42:129-144
- Frewen BE, Chen THH, Howe GT, Davis J, Rohde A, Boerjan W, Bradshaw HD** (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics* 154(2):837-845
- Freyre R, Douches DS** (1994) Development of a model for marker-assisted selection of specific gravity in diploid potato across environments. *Crop Sci.* 34: 1361-1368
- Fridman E, Pleban T, Zamir D** (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proc. Natl. Acad. Sci. USA* 97(9):4718-4723
- Frommer WB and Sonnewald U** (1995) Molecular analysis of carbon partitioning in Solanaceous species. *J. Exp. Bot.* 46:587-607
- Frova C, Gorla MS** (1993) Quantitative expression of maize HSPs: genetic dissection and association with thermotolerance. *Theor. Appl. Genet.* 86:213-220
- Fu H and Park WD** (1995) Sink- and vascular-associated sucrose synthase functions are encoded by different gene classes in potato. *Plant Cell* 7 (9):1369-1385
- Ganal MW, Czhil R, Hannappel U, Kloos DU, Polley A, Ling HQ** (1998) Sequencing of cDNA clones from the genetic map of tomato (*Lycopersicon esculentum*) *Genome Research* 8:842-847
- Ganguly A, Rock MJ, Prockop DJ** (1993) Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double stranded PCR products and DNA fragments: Evidence of solvent-induced bands in DNA heteroduplexes. *Proc. Natl. Acad. Sci. USA* 90:10325-10329
- Gebhardt C, Mugniery D, Ritter E, Salamini F, Bonnel E** (1993) Identification of RFLP markers closely linked to the H1 gene conferring resistance to *Globodera rostochiensis* in potato. *Theor. Appl. Genet.* 85:541-544



## References

---

- Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganai MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homologous tomato genome. *Theor. Appl. Genet.* 83: 49-57
- Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. *Theor. Appl. Genet.* 78: 65-75
- Gebhardt C, Ritter E, Salamini F (1994) RFLP map of potato. In: Phillips RL, Vasil IK (eds) DNA-based markers in plants. Kluwer Academic Publishers, Dordrecht Boston London, pp 271-285
- Gerats T, De Keukeleire P, Deblaere R, Van Montagu M, Zethof J (1995) Amplified fragment length polymorphism (AFLP) mapping in Petunia; a fast and reliable method for obtaining a genetic map. *Act. Hort.* 420:58-61
- Goldman IL, Rocheford TR, Dudley JW (1993) Quantitative trait loci influencing protein and starch concentration in the Illinois long term selection maize strains. *Theor. Appl. Genet.* 87:217-224
- Gould SJ, Subramani S, and Scheffer IE (1989) Use of the DNA polymerase chain reaction for homology probing: Isolation of partial cDNA or genomic clones encoding the iron-sulfur protein of succinate dehydrogenase from several species. *Proc. Natl. Acad. Sci. USA* 86:1934-1938
- Graeve K, von Schaeuwen A and Scheibe R (1994) Purification, characterization, and cDNA sequence of glucose-6-phosphate dehydrogenase from potato (*Solanum tuberosum* L.). *Plant J.* 5 (3):353-361
- Grof CP, Winning BM, Scaysbrook TP, Hill SA and Leaver CJ (1995) Mitochondrial pyruvate dehydrogenase. Molecular cloning of the E1 alpha subunit and expression analysis. *Plant Physiol.* 108 (4):1623-1629
- Groose RW, Kojis WP, Bingham ET (1988) Combining ability differences between isogenic diploid and tetraploid alfalfa. *Crop Sci.* 28:7-10
- Guillaudeux T, Janer M, Wong GK, Spies T, Geraghty DE (1998) The complete genomic sequence of 424015bp at the centromeric end of the HLA class I region: gene content and polymorphism. *Proc. Natl. Acad. Sci. USA* 95:9494-9499
- Haanstra JPW, Wye C, Verbakel H, Meijer-Dekens F, Van den Berg P, Odinet P, Heusden AW, Tanksley S, Lindhout P, Peleman J (1999) An integrated high-density RFLP-AFLP map of tomato based on two *Lycopersicon esculentum* x *L. pennellii* F2 populations. *Theor. Appl. Genet.* 99:254-271
- Hacia JG, Fan JB, Ryder O, Jin L, Edgemon K, Ghandour G, Mayer RA, Sun B, Hsie L, Robbins CM, Brody LC, Wang D, Lander ES, Lipshutz R, Fodor SPA, Collins FS (1999) Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nature Genetics* 22(2):164-167
- Haff LA, Smirnov IP (1997) Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Res.* 7: 378-388
- Halward T, Stalker T, LaRue E, Kochert G (1992). Use of single-primer DNA amplifications in the genetic studies of peanut (*Arachis hypogaea* L.). *Plant Mol. Biol.* 18:325-325
- Hamalainen J, Watanabe K, Valkonen J, Arihara A, Plaisted R, Pehu E, Miller L, Slack S (1997) Mapping and marker assisted selection for a gene for extreme resistance to potato virus-Y. *Theor. Appl. Genet.* 94:192-197
- Harms K, Wohner RV, Schulz B and Frommer WB (1994) Isolation and characterization of Ptype H<sup>+</sup>-ATPase genes from potato. *Plant Mol. Biol.* 26 (3):979-988
- Harris N, Taylor JE and Roberts JA (1994) Isolation of a mRNA encoding a nucleoside diphosphate kinase from tomato that is up-regulated by wounding. *Plant Mol. Biol.* 25 (4):739-742
- Hartings H (1999) High resolution fingerprinting of transcribed genes by means of a modified cDNA-AFLP method. *Maydica* 44(2):179-181

## References

---

- Hartl L, Seefelder S** (1998) Diversity of selected hop cultivars detected by fluorescent AFLPs. *Theor. Appl. Genet.* 96:112-116
- Hawkes JG** (1978) History of the potato. In: *The Potato Crop*, ISBN 0 412 128306, pp8, Chapman and Hall Ltd
- Hawkes JG** (1990) *The Potato, evolution, biodiversity and genetic resources*, Belhaven press, London
- Hedley PE, Machray GC, Davies HV, Burch L and Waugh R** (1994) Potato (*Solanum tuberosum*) invertase-encoding cDNA and their differential expression. *Gene* 145 (2):211-214
- Helentjaris T** (1987) A genetic linkage map for maize based on RFLPs. *Trends Genet.* 3:217-221
- Helentjaris T, Slocum M, Wright S, Schaefer A, Nienhuis J** (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theor. Appl. Genet.* 72:761-769
- Hill SA** (1997) Carbon metabolism in mitochondria. In: *Plant Metabolism* pp181-196, 2<sup>nd</sup> edition, ISBN 0-582-25906-1, Addison Wesley Longman Limited 1997
- Holland PW** (1993) Cloning genes using the polymerase chain reaction. In *Essential developmental biology, a practical approach*. Edited by A. Stern and P.W. Holland. IRL Press, Oxford.
- Horton D** (1988) Potato: truly a world crop. *SPAN* 30:116-118
- Horton R, Niblett D, Milne S, Palmer S, Tubby B, Trowsdale J, Beck S** (1998) Large-scale sequence comparisons reveal unusually high levels of variation in the HLA-DQB1 locus in the class II region of the human MHC. *J. Mol. Biol.* 282:71-97
- Huang J, and Sun M** (1999) A modified AFLP with fluorescence-labelled primers and automated DNA sequencer detection for efficient fingerprinting analysis in plants. *Biotechnology Techniques* 13(4):277-278
- Huang XQ, Zeller FJ, Hsam SLK, Wenzel G, Mohler V** (2000) Chromosomal location of AFLP markers in common wheat utilizing nulli-tetrasomic stocks. *Genome* 43:298-305
- Jacobs JME, Van Eck HJ, Arens P, Verkerk-Bakker B, Te Lintel Hekkert B, Bastiaanssen HJM, El-kharbotly A, Pereira A, Jacobsen E, Stiekema WJ** (1995) A genetic map of potato (*Solanum tuberosum*) integrating molecular markers, including transposons, and classical markers. *Theor. Appl. Genet.* 91: 289-300
- Jacobs JME, Van Eck HJ, Horsman K, Arens P, Verkerk-Bakker B, Jacobsen E, Pereira A, Stiekema WJ** (1996) Mapping of resistance to the potato cyst nematode *Globodera rostochiensis* from the wild potato species *Solanum vernei*. *Mol. Breed.* 2:51-60
- Jin H** (1998) Identification of quantitative loci for tolerance to barley yellow dwarf virus in oat. *Phytopathology* 88:410-415
- Jones CJ** (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol. Breed.* 3:381-390
- Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JDG** (1994) Isolation of the tomato Cf-9 gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266:789-793
- Kammerer B, Fischer K, Hilpert B, Schubert S, Gutensohn M, Weber A and Flugge UI** (1998) Molecular characterization of a carbon transporter in plastids from heterotrophic tissues: the glucose 6-phosphate/phosphate antiporter. *Plant Cell* 10 (1):105-117
- Kammerer B, Fischer K, Hilpert B, Schubert S, Gutensohn M, Weber A and Fluege, U. I.** (1998). Molecular characterization of a carbon transporter in plastids from heterotrophic tissues: the glucose 6-phosphate/phosphate antiporter. *Plant Cell* 10:105-117
- Kaufmann H, Salamini F, and Thompson RD** (1991) Sequence variability and gene structure at the self-incompatibility locus of *Solanum tuberosum*. *Mol. Gen. Genet.* 226:457-466

## References

---

- Kawchuk LM, Lynch DR, Thomas J, Penner B, Sillito D, and Kulcsar F** (1996) Characterization of *Solanum tuberosum* simple sequence repeats and application to potato cultivar identification. *American potato J.* 73:325-335
- Kearsey MJ, Farquhar AGL** (1998) QTL analysis in plants; where are we now? *Heredity* 80:137-142
- Keim P** (1997) Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers. *J. Bacteriol.* 179:818-824
- Khoshnoodi J, Blennow A, Ek B, Rask L and Larsson H** (1996) The multiple forms of starch-branching enzyme I in *Solanum tuberosum*. *Eur. J. Biochem.* 242 (1):148-155
- Kianian SF, Egli MA, Phillips RL, Rines HW, Somers DA, Gengenbach BG, Webster FH, Livingston SM, Groh S, O'Donoghue LS; Sorrells ME, Wesenberg DM, Stuthman DD, Fulcher RG** (1999) Association of a major groat oil content QTL and an acetyl-CoA carboxylase gene in oat. *Theor. Appl. Genet.* 98:884-894
- Knapp SJ, Bridges WC** (1990) Using molecular markers to estimate quantitative trait locus parameters; power and genetic variances for unreplicated and replicated progeny. *Genetics* 126:769-777
- Kombrink E, Somssich LE** (1997) Pathogenesis related proteins and plant defense. In: *The Mycota V Part A*, Carroll/Tudzynski (Eds), Springer-Verlag Berlin Heidelberg, pp 107 – 128.
- Konieczny A, Ausubel FM** (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *The Plant Journal* 4:403-410
- Kosambi DD** (1944) The estimation of map distances from recombination values. *Ann Eugen* 12:172-175
- Kossmann J, Mueller-Roeber B, Dyer TA, Raines CA, Sonnewald U and Willmitzer L** (1992) Cloning and expression analysis of the plastidic fructose-1, 6-bisphosphatase coding sequence from potato: circumstantial evidence for the import of hexoses into chloroplasts. *Planta* 188: 7-12
- Kreitman M, Aquade M** (1986) Genetic uniformity in two populations of *Drosophila melanogaster* as revealed by four-cutter filter hybridisation. *Proc. Natl. Acad. Sci. USA* 83:3562-3566
- Kresovich S, Szewze-McFadden AK, Bliet SM, McFerson JR** (1995) Abundance and characterization of simple sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. *Theor. Appl. Genet.* 91:206-211
- Kruger NJ** (1997) Carbohydrate synthesis and degradation. In: *Plant Metabolism* pp83-102, 2<sup>nd</sup> edition, ISBN 0-582-25906-1, Addison Wesley Longman Limited 1997
- Kurata N, Nagamura Y, Yamamoto K, Harushima Y, Sue N, Wu J, Antonio BA, Shomura A, Shimizu T, Lin S-Y, Inoue T, Fukuda A, Shimano T, Kuboki Y, Toyama T, Miyamoto Y, Kiriwara T, Hayasaka K, Miyao A, Monna L, Zhong HS, Tamura Y, Wang Z-X, Momma T, Umehara Y, Yano M, Sasaki T, Minobe Y** (1994) A 300 kilobase interval genetic map of rice including 883 expressed sequences. *Nature Genetics* 8: 365-372
- Lagercrantz U** (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics* 150(3):1217-1228
- Lagercrantz U, Putterill J, Coupland G, Lydiate D** (1996) comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. *Plant Journal* 9(1):13-20
- Laloi M, Klein M, Riesmeier JW, Muller-Rober B, Fleury C, Bouillaud F and Ricquier D** (1997) A plant cold-induced uncoupling protein. *Nature* 389 (6647):135-136
- Lander ES, and Botstein D** (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199

## References

---

- Landry BS, Kesseli RV, Farrara B, Micheltmore RW** (1987b) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. *Genetics* 116:331-337
- Landschuetze V, Mueller-Roeber B, and Willmitzer L** (1995a) Mitochondrial citrate synthase from potato: predominant expression in mature leaves and young flower buds. *Planta* 196:756-764.
- Landschuetze V, Willmitzer L, and Mueller-Roeber B** (1995b) Inhibition of flower formation by antisense repression of mitochondrial citrate synthase in transgenic potato plants leads to specific disintegration of the ovary tissues of flowers. *EMBO J.* 14:660-66
- Larrick, JW, Danielson L, Brenner CA, Wallace EF, Abrahamson M, Fry KE, and Borrebaeck CA** (1989) Polymerase chain reaction using mixed primers: Cloning of human monoclonal antibody variable regions genes from single hybridoma cells. *Bio/Technology* 7:934-938
- Larsson CT, Khoshnoodi J, Ek B, Rask L and Larsson H** (1998) Molecular cloning and characterization of starch-branching enzyme II from potato. *Plant Mol. Biol.* 37 (3):505-511
- Lefebvre DD, Gellatly KS** (1997) Fundamentals of gene structure and control. In: *Plant Metabolism* p8 2<sup>nd</sup> edition, Addison Wesley Longman Limited, Singapore
- Leggiewie G, Willmitzer L and Riesmeier JW** (1997) Two cDNAs from potato are able to complement a phosphate uptake-deficient yeast mutant: identification of phosphate transporters from higher plants. *Plant Cell* 9 (3):381-392
- Leonards-Schippers C, Gieffers W, Schaefer-Pregl R, Ritter E, Knapp SJ, Salamini F, Gebhardt C** (1994) Quantitative resistance to *Phytophthora infestans* in potato: a case study for QTL mapping in an allogamous plant species. *Genetics* 137:67-77
- Li W, Sadler LA** (1991) Low nucleotide diversity in man. *Genetics* 129:513-523
- Li X, Van Eck HJ, Rouppe van der Voort JNAM, Huigen DJ, Stam P, Jacobsen E** (1998) Autotetraploid and genetic mapping using common AFLP markers: the R2 allele conferring resistance to *Phytophthora infestans* mapped on potato chromosome 4. *Theor. Appl. Genet.* 96:1121-1128
- Lin X, Kaul S, Rounsley S, Benito MI, Town CD, Fujii CY, Mason T, Bowman CL, Barnstead M, Feldblyum TV, Buell CR, Ketchum KA, Lee J, Ronning CM, Koo HL, Moffat KS, Cronin LA, Shen M, Pai G, van Aken S, Umayan L, Tallon LJ, Gill JE, Adams MD, Carrera AJ, Creasy TH, Goodman HM, Somerville CR, Copenhaver GP, Preuss D, Nierman WC, White O, Eisen JA, Salzberg SL, Fraser CM, Venter C** (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* 402: 761-768
- Lindblad-Toh K, Winchester E, Daly MJ, Wang DG, Hirschhorn JN, Laviolette JP, Ardlie K, Reich DE, Robinson E, Sklar P, Shah N, Tomas D, Fan JB, Gingeras T, Warrington J, Patil N, Hudson TJ, Lander ES** (2000) Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nature Genetics* 24:381-386
- Lister C, and Dean C** (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant J.* 4:745-750
- Liu Z** (1998) Inheritance and usefulness of AFLP markers in channel catfish (*Ictalurus punctatus*), blue catfish (*I. furcatus*), and their F1, F2 and backcross hybrids. *Mol. Gen. Genet.* 258:260-268
- Lu ZX** (1998) Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks. *Genome* 41:199-207
- Lunden AP** (1956) Undersøkelser over forholdet mellom potetenes spesifikke vekt og deres tørrstoff- og stivelsesinnhold. *Forskn. Forsøk Landbruket* 7:81-107
- Lynch M, and Milligan BG** (1994) Analysis of population genetic structure with RAPD markers. *Mol. Ecol.* 3:91-99

## References

---

- Maheswaran M, Subudhi PK, Nandi S, Xu JC, Parco A, Yang DC, Huang N** (1997) Polymorphism, distribution, and segregation of AFLP markers in a doubled-haploid rice population. *Theor. Appl. Genet.* 94:39-45
- Maldonado LA, Wright JE and Scott GJ** (1998) Constraints to production and use of potato in Asia. *Amer. J. of Potato Res.* 75:71-79
- Martin G, Williams JGK, Tanksley SD** (1991) RAPD identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and nearly isogenic lines. *Proc. Natl. Acad. Sci. USA* 88:2236-2340
- Mather K** (1949) *Biometrical Genetics*. 1<sup>st</sup> Edn, Methuen, London
- Maughan PJ, Saghai Maroof MA, Buss GR** (1995) Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. *Genome* 38:715-723
- McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman, WR, Tanksley SD** (1988) Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* 76:815-829
- McMullen MD, Byrne PF, Snooks ME, Wiseman BR, Lee EA, Widstrom NW, Coe EH** (1998) Quantitative trait loci and metabolic pathways. *Proc. Natl. Acad. Sci. USA* 95:1996-2000
- McPherson MJ, Jones KJ, and Gurr SJ** (1991) PCR with highly degenerate primers. *In* PCR, a practical approach. *Edited by* M.C. McPherson, P. Quirke, G.R. Taylor. IRL Press, Oxford.
- Meksem K, Leister D, Peleman J, Zabeau M, Salamini F, Gebhardt C** (1995) A high-resolution map of the vicinity of the R1 locus on chromosome V of potato based on RFLP and AFLP markers. *Mol. Gen. Genet.* 249:74-81
- Merkelbach S, Gehlen J, Denecke M, Hirsch HJ and Kreuzaler F** (1993) Cloning, sequence analysis and expression of a cDNA encoding active phosphoenolpyruvate carboxylase of the C3 plant *Solanum tuberosum*. *Plant Mol. Biol.* 23 (4):881-888
- Meyer RC, Hedley PE, Waugh R and Machray GC** (1996) Organisation and expression of a potato (*Solanum tuberosum*) protein kinase gene. *Plant Sci.* 118:71-80
- Meyer RC, Milbourne D, Hackett CA, Bradshaw JE, McNichol JW, Waugh R** (1998) Linkage analysis in tetraploid potato and association of markers with quantitative resistance to late blight (*Phytophthora infestans*) *Mol. Gen. Genet.* 259(2):150-160
- Michelmore RW, Paran I, and Kesseli RV** (1991) Identification of markers linked to disease resistance genes by bulked segregation analysis: a rapid method to detect markers in specific genomic regions by using segregating populations *Proc. Natl. Acad. Sci. USA* 88:9828-9832
- Miller ME, and Chourey PS** (1992). The maize invertase-deficient *miniature-1* seed mutation is associated with aberrant pedicel and endosperm development. *Plant Cell* 4:297-305
- Mindrinis M, Katagari F, Yu GL, Ausubel FM** (1994) The *A. thaliana* disease resistance gene RPS2 encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* 78:1089-1099
- Moehs CP, Allen PV, Friedman M, and Belknap WR** (1996). Cloning and expression of transaldolase from potato. *Plant Mol. Biol.* 32: 447-452
- Morgante M and Olivieri AM** (1993) PCR-based microsatellites as markers in plant genetics. *Plant J.* 3:175-182
- Mori H, Tanizawa K, Fukui T** (1991) Potato tuber type H phosphorylase isozyme. Molecular cloning, nucleotide sequence, and expression of a full-length cDNA in *Escherichia coli*. *J. Biol. Chem.* 266:18446-18453
- Mueller UG, Wolfenbarger LL** (1999) AFLP genotyping and finger printing. *TREE* 14:38-394

## References

---

- Muller-Rober BT., Kossmann J, Hannah LC, Willmitzer L. And Sonnewald U** (1990) One of two different ADP-glucose pyrophosphorylase genes from potato responds strongly to elevated levels of sucrose. *Mol. Gen. Genet.* 224 (1):136-146
- Nakano K, Mori H and Fukui T** (1989) Molecular cloning of cDNA encoding potato amyloplast alpha-glucan phosphorylase and the structure of its transit peptide. *J. Biochem.* 106 (4):691-695
- Nakata PA, Greene TW, Anderson JM, Smith-White BJ, Okita TW and Preiss J** (1991) Comparison of the primary sequences of two potato tuber ADP-glucose pyrophosphorylase subunits. *Plant Mol. Biol.* 17 (5):1089-1093
- Neuffer MG, Coe EH, Wessler SR** (1997) Mutants of maize. Cold Spring Harbor Laboratory Press.
- Nickerson DA, Tobe VO, and Taylor SL** (1997) PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nuc. Acid Res.* 25:2645-2751
- Nolte KD, and Koch KE** (1993) Companion-cell specific localization of sucrose synthase in zone of phloem loading and unloading. *Plant Physiol.* 101:899-905
- Oberhagemann P, Chatot-Balandras C, Schaefer-Pregl R, Weger D, Palomino C, Salamini F, Bonnel E, Gebhardt C** (1999) A genetic analysis of quantitative resistance to late blight in potato: towards marker-assisted selection. *Molecular Breeding* 5:399-415
- Ortiz R** (1998) Potato breeding via ploidy manipulations. In: *Plant Breeding Reviews* Vol. 16, pp16-19 edited by Jules Janick, ISBN 0-471-25446-0, John Wiley and Sons, Inc.
- Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T** (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA* 86:2766-2700
- Pakniyat H** (1997) AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography. *Genome* 40:332-341
- Patterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD** (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335:721-726
- Pear JR, Sanders RA, Summerfelt KR, Martineau B and Hiatt WR** (1993) Simultaneous inhibition of two tomato fruit cell wall hydrolases, pectinmethylesterase and polygalacturonase, with antisense gene constructs. *Antisense Res. Dev.* 3 (2):181-190
- Penner GA, Bush A, Wise R, Kim W, Domier L, Kasha K, Laroche A, Scoles G, Molnar SJ, Fedak G** (1993) Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. *PCR Methods Appl.* 2:341-345
- Pflieger S, Lefebvre V, Caranta C, Blattes A, Goffinet B, Palloix A** (1999) Disease resistance gene analogs as candidates for QTLs involved in pepper-pathogen interactions. *Genome* 42(6):1100-1110
- Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce-Jacino M** (1999) Mining SNPs from EST databases. *Genome Res.* 9:167-174
- Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A** (1996b) The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.* 2:225-238
- Qi X, Stam P, Lindhout P** (1998) Use of locus-specific AFLP markers to construct a high-density molecular map in barley. *Theor. Appl. Genet.* 96:376-384
- Quarrie SA** (1997) QTL analysis to study the association between leaf size and abscisic acid accumulation in droughted rice leaves and comparisons across cereals. *Plant Mol. Biol.* 35:155-165
- Queller DC, Strassmann JE, and Hughes CE** (1993) Microsatellites and kinship. *Trends Ecol. Evol.* 8:285-288

## References

---

- Quick WP, Neuhaus HE, Sonnewald U, Preiss J** (1997) In: A molecular approach to primary metabolism in higher plants pp41-104, Taylor and Francis Ltd, London
- Reamon-Buettner SM, Jung C** (2000) AFLP-derived STS markers for the identification of sex in *Asparagus officinalis* L. Theor. Appl. Genet. 100:432-438
- Ridout CJ, Donini P** (1999) Use of AFLP in cereals research. Trend in Plant Science 4:76-79
- Riesmeier JW, Hirner B and Frommer WB** (1993) Potato sucrose transporter expression in minor veins indicates a role in phloem loading. Plant Cell 5 (11):1591-1598
- Ritter E, Debener T, Barone A, Salamini F, Gebhardt C** (1991) RFLP mapping on potato chromosomes of two genes controlling extreme resistance to potato virus X (PVX). Mol. Gen. Genet. 227:81-85
- Ritter E, Gebhardt C, Salamini F** (1990) Estimation of recombination frequencies and construction of RFLP linkage maps in plant from crosses between heterozygous parents. Genetics 125:645-654
- Ritter E, Salamini F** (1991) The calculation of recombination frequencies in crosses of allogamous plant species with applications to linkage mapping. Genet. Res. 67:55-67
- Rohde W, Becker D, Kull B and Salamini F** (1990) Structural and functional analysis of two waxy gene promoters from potato. J. Gen. and Breed. 44:311-315
- Ronfort J, Jenczewski E, Bataillon T, and Rousset F** (1998) Analysis of population structure in autotetraploid species. Genetics 150:921-930
- Rongwen J, Akkaya MS, Lavi U, Cregan PB** (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor. Appl. Genet. 90:43-48
- Rosendahl S, Taylor JW** (1997) Development of multiple genetic markers for studies of genetic variation in arbuscular mycorrhizal fungi using AFLP. Mol. Ecol. 6:821-829
- Roupe van der Voort JNAM, Van Eck HJ, Folkertsma RT, Van Zandvoort P, Hutten RCB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J** (1997) Locus specificity of AFLP markers used to align genetic maps from different potato genotypes. Mol. Gen. Genet. 255:438-447
- Roupe van der Voort J, Wolters P, Folkertsma R, Hutten R, van Zandvoort P, Vinke H, Kanyuka K, Bendahmane A, Jacobsen E, Janssen R, Bakker J** (1997) Mapping of the cyst nematode resistance locus *Gpa2* in potato using a strategy based on comigrating AFLP markers. Theor. Appl. Genet. 95:874-880
- Roupe van der voort JNAM, Vaneck HJ, Draaistra J, Vanzandvoort PM, Jacobsen E, Bakker J** (1998) An online catalogue of AFLP markers covering the potato genome. Mol. Breeding 4(1):73-77
- Russell JR** (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor. Appl. Genet. 95:714-722
- Saghai Maroof MA, Biyashev RM, Yang GP, Zhang Q, and Allard RW** (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc. Natl. Acad. Sci. USA 91:5466-5470
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA** (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491
- SAS Institute Inc** (1990) SAS Language: reference, version 6, fourth edition, vol. 2, SAS Institute Inc., Cary, North Carolina
- Sasaki T, Song J, Koga-Ban Y, Matsui E, Fang F, Higo H, Nagasaki H, Hori M, Miya M, Murayama-Kayano, Takiguchi T, Takasuga A, Niki T, Ishimaru K, Ikeda H, Yamamoto Y, Mukai Y, Ohta I, Miyadera N, Havukkala I, and Minobe Y** (1994) Towards cataloguing all rice genes: Large scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. Plant J. 6:615-624
- Sax K** (1923) The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552-560

## References

---

- Schäfer-Pregl R, Ritter E, Concilio L, Hesselbach J, Lovatti L, Walkemeier B, Thelen H, Salamini F, Gebhardt C** (1998) Analysis of quantitative trait loci (QTLs) and quantitative trait alleles (QTAs) for potato tuber yield and starch content. *Theor. Appl. Genet.* 97:834-846
- Schneider A** (1995) Expression und Regulation des durch Kälte induzierbaren Gens *ci21* aus der Kartoffel. PhD Thesis, University of Cologne, Germany.
- Schneider K, Borchardt DC, Schafer-Pregl R, Nagl N, Glass C, Jeppsson A, Gebhardt C, Salamini F** (1999) PCR-based cloning and segregation analysis of functional gene homologues in *Beta vulgaris*. *Mol. Gen. Genet.* 262(3):515-524
- Schneider K, Bernd Weisshaar, Diethrich C, Borchardt DC, Salamini F** (2000) Frequency of EST alleles in sugar beet breeding lines based on SNP scanning. *Mol. Breed.* (in press)
- Schneider K, Douches DS.** (1997) Assessment of PCR-based simple sequence repeats to fingerprint north American potato cultivars. *American Potato Journal* 74(3):149-160
- Schondelmaier J, Steinruecken G, Jung C** (1996) Integration of AFLP markers into a linkage map of sugar beet (*Beta vulgaris* L.). *Plant Breeding* 115:231-237
- Schut JW, Qi X, and Stam P** (1997) Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. *Theor. Appl. Genet.* 95:1161-1168
- Schwarz G, Michalek W, Mohler V, Wenzel G, Jahoor A** (1999) Chromosome landing at the *Mla* locus in barley (*Hordeum vulgare* L.) by means of high-resolution mapping with AFLP markers. *Theor. Appl. Genet.* 98:521-530
- Semblat JP** (1998) High-resolution DNA fingerprinting of parthenogenetic root-knot nematodes using AFLP analysis. *Mol. Ecol.* 7:119-125
- Senior ML, Chin ECL, Lee M, Smith JSC, Stuber CW** (1996) Simple sequence repeat markers developed from maize sequences found in the Genbank database: map construction. *Crop Sci.* 36:1676-1683
- Shan X, Blake TK, Talbert LE** (1999) Conversion of AFLP markers to sequence-specific PCR markers in barley and wheat. *Theor. Appl. Genet.* 98:1072-1078
- Sheffield VC, Beck JS, Kwitek AE, Sandstrom DW, Stone EM** (1993) The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics* 16:325-332
- Simko I, Mc Murray, Yang H, Manschot A, Davies P, Ewing E** (1997) Evidence from polygene mapping for a causal relationship between potato tuber dormancy and abscisic acid content. *Plant Physiol.* 115:1453-1459
- Smith SB, Taylor MA, Burch LR and Davies HV** (1993) Primary structure and characterization of a cDNA clone of fructokinase from potato (*Solanum tuberosum* L. cv record). *Plant Physiol.* 102 (3):1043-1043
- Sonnenwald U, Basner A, Greve B and Steup M** (1995) A second L-type isozyme of potato glucan phosphorylase: cloning, antisense inhibition and expression analysis. *Plant Mol. Biol.* 27 (3):567-576
- Speirs J, Hinde R, Paulin R and Lee E** (1995) Nucleotide sequence of a cDNA encoding tomato glycolate oxidase (Accession No. X92888). *Plant Physiol.* 110:336-336
- STATISTICA™ for Windows (volume I): General Conventions and Statistics.** 2<sup>nd</sup> edition, StatSoft, Inc. 1995.
- Syvanen AC** (1999) From gels to chips: "Minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. *Human Mutation.* 13(1):1-10
- Takaha T, Yanase M, Okada S and Smith SM** (1993) Disproportionating enzyme of potato. Purification, molecular cloning, and potential role in starch metabolism. *J. Biol. Chem.* 268 (2):1391-1396
- Talarczyk A and Hennig J** (1998) Characterization of a cDNA encoding a glucan endo-1,3-beta-glucosidase from potato (*Solanum tuberosum* L.). *Plant Physiol.* 118:712-712



## References

---

- Tanksley SD** (1993) Mapping polygenes. *Ann. Rev. Genet.* 27:205-233
- Tanksley SD, Bernatzky R, Lapitan NL, Prince JP** (1988) Conservation of gene repertoire but not gene order in pepper and tomato. *Proc. Natl. Acad. Sci. USA* 85:6419-6423
- Tanksley SD, Ganai MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND** (1992) High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132:1141-1160.
- Tarn TR, Tai GCC, De Jong H, Murphy AM, Seabrook JEA** (1992) Breeding potatoes for long-day temperate climates. In *Plant Breeding Review Vol. 9*, edited by Jules Janick, ISBN 0-471-57498-8, pp252
- Tautz D** (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* 17:6463-6471
- Taylor MA, George LA, Ross HA and Davies HV** (1998) cDNA cloning and characterisation of an alpha-glucosidase gene from potato (*Solanum tuberosum* L.)
- Teige M, Kopriva S, Bauwe H and Suess KH** (1995) Primary Structure of Chloroplast Transketolase from Potato (Accession No. Z50099)(PGR96-121)
- Teige M, Kopriva S, Bauwe H and Suss KH** (1995) Chloroplast pentose-5-phosphate 3-epimerase from potato: cloning, cDNA sequence, and tissue-specific enzyme accumulation. *FEBS Lett.* 377 (3):349-352
- Thoday JM** (1961) Location of polygenes. *Nature* 191:368-370
- Thomas MR and Scott NS** (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence tagged sites (STSs). *Theor. Appl. Genet.* 86:985-990
- Thornton RE and Sieczka JB** (1980) Commercial potato production in North America. *Amer. Pot. J. Supp.* Vol. 57, p36
- Tingey S, Del Tufo JP** (1993) Genetic analysis with random polymorphic DNA markers. *Plant Physiol.* 101:349-352
- Tohme J, Gonzalez DO, Beebe S, Duque MC** (1996) AFLP analysis of gene pools of a wild bean core collection. *Crop Sci.* 36:1375-1384
- Tragoorung S, Kanazin V, Hayes PM, and Blake TK** (1992) Sequence-tagged-site-facilitated PCR for barley genome mapping. *Theor. Appl. Genet.* 84:1002-1008
- Trimming BA, and Emes MJ** (1993). Glycolytic enzymes in non-photosynthetic plastids of pea (*Pisum sativum* L.) roots. *Planta* 190:439-45
- Tyagi S, Bratu DP, Kramer FR** (1998) Multicolour molecular beacons for allele discrimination. *Nat. Biotechnol.* 16:49-53
- Tyagi S, Kramer FR** (1995) Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* 14:303-308
- Van Berkel J, Salamini F, Gebhardt C** (1994) Transcripts accumulating during cold storage of potato (*Solanum tuberosum* L.) tubers are sequence related to stress-responsive genes. *Plant Physiol.* 104: 445-452
- Van Campenhout S, Koebner RMD, Volckaert G** (2000) The applicability of consensus PCR primers across species and genera: the use of wheat Em sequences to develop markers for orthologues in rye. *Theor. Appl. Genet.* 100:328-336
- Van den Berg JH, Ewing EE, Plaisted RL, McMurry S, Bonierbale MW** (1996a) QTL analysis of potato tuberization. *Theor. Appl. Genet.* 93: 307-316
- Van den Berg JH, Ewing EE, Plaisted RL, McMurry S, Bonierbale MW** (1996b) QTL analysis of potato tuber dormancy. *Theor. Appl. Genet.* 93: 317-324

## References

---

- Van Der Lee T** (1997) AFLP linkage map of the oomycete *Phytophthora infestans*. Fungal Genet. Biol. 21:278-291
- Van der Straeten D, Rodrigues-Pousada RA, Goodman HM and Van Montagu M** (1991) Plant enolase: gene structure, expression, and evolution. Plant Cell 3 (7):719-735
- Van Eck HJ, Jacobs JME, Stam P, Ton J, Stiekema WJ, Jacobsen E** (1994a) Multiple alleles for tuber shape in diploid potato detected by qualitative and quantitative genetic analysis using RFLPs. Genetics 137:303-309
- Van Eck HJ, Jacobs JME, Van den Berg P, Stiekema WJ, Jacobsen E** (1994b) The inheritance of anthocyanin pigmentation in potato (*Solanum tuberosum* L.) and mapping of tuber skin colour loci using RFLPs. Heredity 73:410-421
- Van Eck HJ, Rouppe van der Voort J, Draaistra J, Van Zandvoort P, Van Enckefort E, Segers B, Peleman J, Jacobsen E, Helder J, Bakker J** (1995) The inheritance and chromosomal localisation of AFLP markers in a non-inbred potato offspring. Mol. Breed. 1:397-410
- Van Heusden AW, Koornneef M, Voorrips RE, Bruggemann W, Pet G, R. Vrielink-van Ginkel, Chen Xinwei and Lindhout P** (1999) Three QTLs from *Lycopersicon peruvianum* confer a high level of resistance to *Clavibacter michiganensis* ssp *michiganensis*. Theor. Appl. Genet. 99(6):1068-1074
- Voorrips RE, Jongerius MC, and Kanne HJ** (1997) Mapping of two genes for resistance to clubroot (*Plasmodiophora brassicae*) in a population of doubled haploid lines of *Brassica oleracea* by means of RFLP and AFLP markers. Theor. Appl. Genet. 94:75-82
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M** (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414
- Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen NP, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES** (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280(5366):1077-1082
- Waugh P, Powell W** (1992) Using RAPD markers for crop improvement. Trends Biotech. 10:186-191
- Waugh R, Bonar N, Baird E, Thomas B, Graner A, Hayes P, Powell W** (1997) Homology of AFLP products in three mapping populations of barley. Mol. Gen. Genet. 255:311-321
- Weber JL and May PE** (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet. 44:388-396
- Weining S, and Langrises P** (1991) Identification and mapping of polymorphisms in cereals based on the polymerase chain reaction. Theor. Appl. Genet. 82:209-216
- Welsh J, McClelland M** (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acid Res. 18:7213-7218
- Werr W, Frommer WB, Maas C, Starlinger P** (1985) Structure of the sucrose synthase gene on chromosome 9 of *Zea mays* L. EMBO J. 4: 1373-1380.
- White MB, Carvalho M, Derse D, O'Brien SJ, Dean M** (1992) Detecting single base substitutions as heteroduplex polymorphisms. Genomics 12:301-305
- Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B** (1994) The product of the tobacco mosaic virus resistance gene *N*: similarity to Toll and interleukine-1 receptor. Cell 78:1101-1115
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV** (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acid Res. 18:6531-6535

## References

---

- Winning BM, Bourguignon J and Leaver CJ** (1994) Plant mitochondrial NAD<sup>+</sup>-dependent malic enzyme. cDNA cloning, deduced primary structure of the 59- and 62-kDa subunits, import, gene complexity and expression analysis. *J. Biol. Chem.* 269 (7):4780-4786
- Wolter FP, Fritz CC, Willmitzer L, Schell J** (1988) *rbcS* genes in *Solanum tuberosum*: conservation of transit peptide and exon shuffling during evolution. *Proc. Natl. Acad. Sci. USA* 85: 846-850
- Wright S** (1934) An analysis of variability in number of difits in an inbred strain of guinea pigs. *Genetics* 19:506-536
- Yin X, Stam P, Dourleijn CJ, Kropff MJ** (1999) AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. *Theor. Appl. Genet.* 99(1-2):244-253
- Zabeau M, and Vos P** (1993) Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Application number 92402629-7, publication number 0 534 858 A1
- Zamir D, Tanksley SD** (1988) Tomato genome is comprised largely of fast-evolving, low copy-number sequences. *Mol. Gen. Genet.* 213:254-261
- Zhang LH, Ozias-Akins P, Kochert G, Kresovich S, Dean R, Hanna W** (1999) Differentiation of bermudagrass (*Cynodon* spp.) genotypes by AFLP analyses. *Theor. Appl. Genet.* 98:895-902
- Zhao J and Herrmann KM** (1993) Cloning and sequence of a second cDNA encoding 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from *Solanum tuberosum*. *Plant Physiol.* 100:1075-1076
- Zrenner R, Krause KP, Apel P and Sonnewald U** (1996) Reduction of the cytosolic fructose-1, 6-bisphosphatase in transgenic potato plants limits photosynthetic sucrose biosynthesis with no impact on plant growth and tuber yield. *Plant J.* 9 (5):671-681
- Zrenner R, Salanoubat M, Willmitzer L and Sonnewald U** (1995) Evidence of the crucial role of sucrose synthase for sink strength using transgenic potato plants (*Solanum tuberosum*). *Plant J.* 7 (1):97-107
- Zrenner R, Schuler K and Sonnewald U** (1996) Soluble acid invertase determines the hexose-to-sucrose ratio in cold-stored potato tubers. *Planta* 198 (2):246-252
- Zubay GL** (1998) Metabolism of carbohydrate. In: *Biochemistry* 4<sup>th</sup> edition, Wm. C. Brown Publishers, pp265-438

## 7. APPENDIX

### Six new nucleotide sequences of functional genes in potato\*

#### 7.1 Partial potato genomic sequence of gene *Rca*<sup>a</sup>

1	<u>ACACCGTCAA</u>	<u>CAACCAGATG</u>	GTGAATGCCA	CCCTCATGAA	CATTGCTGAC
51	AACCCAAACAA	ATGTCCAGCT	CCCCGGTATG	TACAACAAGC	AAGAGAACTG
101	CAGGGTCCCC	ATTATTGTCA	CTGGTAACGA	CTTCTCCACA	TTGTATGCTC
151	CTCTTATCCG	TGATGGTCGT	ATGGAGAAAGT	TCTACTGGGC	ACCAACTAGG
201	GAGGATAGAA	TTGGTGTTTG	CAAGGGTATT	TTCAGAACTG	ACAACGTGCC
251	TGAGGAAGCT	GTTGTAAAGA	TTGTCGATTC	CTTCCCTGGA	CAATCTATTG
301	GTACAAACAC	TAAAGAATTC	AAAATGAAAT	TTCTCTTTTA	A

#### 7.2 Partial potato genomic sequence of gene *Ant*<sup>a</sup> *Sequence I*

1	<u>TGGAGAGGAA</u>	<u>ACACTGCCAA</u>	<u>TGTTATCCGT</u>	TATTTCCCCA	CTCAGGTTTA
50	CCATTTGTCT	TCATTTTGAG	GGCTGTAATT	TATGCATCAA	AGATACATGC
101	ACATTTATTG	CTCTGTTGTT	TGCTTAAGTC	GAGCATTGAC	<b>NGACTGATGC</b> (C/T)
151	TCTGTTCCAC	AGGCCCTGAA	CTTTGCATTC	AAGGACTACT	TCAAGAGACT
201	CTTCAACTTC	AAGAAGGACC	GTGATGGCTA	CTGGAAGTGG	TTTGCTGGCA
251	ACCTTGCCCTC	AGGTGGTGCT	GCTGGTGCTT	CTTCTTTGTT	CTTTGTCTAC
301	TCCTTGGAAT	ATGCTCGTAC	CCGTCTTGCT	AATGACGCCA	AGGCTTCAAA
351	GAAGGGAGGT	GAGAGGCAGT	TCAATGGTTT	GGTTGATGTC	TACAAGAAGA
401	CACTTAAATC	TGATGGAATT	GCTGGTCTAT	ACCGTGGATT	CAACATTTCA
451	TGTGTTGGTA	TCATTGTTTA	CCGTGGTTT <b>N</b> (A/G)	TACTTTGGAA	TGTACGACTC
501	CTTGAAGCCT	GTCCTCTTGA	CAGGAAACCT	GCAGGTTAGT	GTTTTTATAA
551	TGTTGGCTTT	GCTATGCTTG	TTCA		

*Sequence II*

```

1  CCTGGGTTGT TCCTGTTCCA TTTTTTTAAA GTATGGAATG CTGGTTAATA
51  TTGTGTTCCA CAGGCATTGA ACTTTGCATT TAAGGACTAC TTCAAGAGGC
101 TCTTCAACTT CAAGAAGGAC CGCGATGGCT ACTGGAAGTG GTTTGCAGGT
151 AACCTGGCAT CTGGTGGTGG TGCTGGTGCT TCCTCTTTGC TCTTTGTTTA
201 CTCCCTTGAC TATGCTCGTA CTCGTCTTGC AAATGATGCC AAGGCTGCAA
251 AGAAGGGAGG TGGTGGCAGA CAATTCGATG GGTTGGTTGA TGTCTACAGA
301 AAGACACTTA AATCTGATGG AGTTGCTGGC TTGTACCGTG GGTTTAAACAT
351 TTCATGTGTT GGTATCATTG TGTACCGTGG TTTGTACTTT GGAATGTATG
401 ATTCATTGAA GCCAGTGCTG TTGACTGGAA AGATGGAGGT TAGTTATCCA
451 TCTATTTTTG GTATTTTCCC ATATTGCTAG TTTGCTATTT TGTGATGATA
501 TCATATGTTG ATAGTTATGT TCAGATGCTT ACAATACTTG GAGATTGTTG
551 CCTGTTACTT TGGAGTTGGG TTTTAGGATG TTATGGTTGG AATATGACTT
601 CATTCTATTT TTTTATTNTC TTGATGGTTG CAGGATAGTT TCTTTGCTAG
      (C/T)
651 CTTTGCTCTT GGATGGCTCA TCACCAATGG TGCTGGTCTT GCATCGTACC
701 CTATTGACAC AGTTAGGAGA AGAATGATGA TGACATCTGG TGAGGCAGTG
751 AAGTACAAGA GCTCGTTCGA TGCATTCAAC CAAATCCTTA AGAATGAGGG
801 TCCCAAATCA CTCTTCAAGG GAGCTGGTGC CAACAT

```

**7.3 Partial potato genomic sequence of gene *Ndps*<sup>a</sup>**

```

1  ATGATCAAGC CTGATGGTGT CCAGTGGCCT GGTGTTGAG ATTATCGGCA
51  GATTTGAGAA GAAAGGATTT TCTTTGAAAAG GCTTGAAGCT CATCACTGTG
101 AACCATGCCT TTGCTGAGAA GCATTACGCT GACTTGTCTG CTAAGCCTTT
151 CTTTAATGGG CTTGTTGAGT ATATTGTTTC TGGACCTGTT GTTGCTATGG
201 TCTGGGAGGG TAAGGGTGTA GTTGCCACTG GCAGGAAGAT CATTGGAGCA
251 ACCAACCCCT TGGAGTCGGC TGCTAGTACC ATCCGTGGTG ATTTTGCTAT
301 TGATATTGGC AGGAATGTTA TTCATAGAAG TGATGTTGTT GAGAGTGCTA
351 GGAAGGAGAT TGCTTTTTTG TCCCCGAAG GAATTGCAGA GTGGCAGAGC
401 AGCCTTCACT CTT

```

## 7.4 Partial potato genomic sequence of gene *Eno*<sup>b</sup>

1	AGCTGTTGTT	AGATAATCTG	GTGGCTGAGA	TGATTTTCTG	TTTAGGTGGT
51	CATTGGAATG	GATGTTGCAG	CATCTGAATT	TTACGGAAAG	GACAAATCTT
101	ATGACCTGAA	CTTCAAAGAA	GAGGCAAGTA	<i>CTGACTATTG</i>	<i>CTTCTTTTAC</i>
151	<i>ACATGC'TTTT</i>	<i>CCGAGGTAGT</i>	<i>CTCGTAGTCA</i>	<i>AACTGTTTAT</i>	<i>ACGTTATTCA</i>
201	<i>GTTTTGCTCA</i>	<i>TATTTATTCT</i>	<i>TATTTGCAGA</i>	GCAATGACGG	CTCACAAAAG
251	ATATCAGGTG	ACCAACTCAA	GGATTTGTAC	AAGTCATTTG	TGTCCGAGTA
301	CCCTATTGTT	TCAATTGAAG	ATCCATTTGA	CCAAGATGAC	TGGGAGACCT
351	ATGCTAAGCT	CACCACTGAG	ATTGGGGAGC	AAGTACAGAT	TGTCGGAGAT
401	GATCTCCTTG	TCACCAACCC	TAAGGTAAGA	<i>TTGAAGCTTT</i>	<i>AACAATATTG</i>
451	<i>CTTGCAATATC</i>	<i>AGAGTTGCTG</i>	<i>TTTCCATTGA</i>	<i>AAAGAAAAGC</i>	<i>AGACATCAAT</i>
501	<i>TTCTCTAGGT</i>	<i>GTTGATGTTT</i>	<i>ACCATGTATA</i>	<i>CTTATTTATA</i>	<i>ATGGTGTGCA</i>
551	GAGGGTCGCC	AAGGCAATTG	CAGAGAAGAC	TTGCAATGCT	CTTCTTCTCA
601	AGGTATCATC	<i>CTATCCAGCT</i>	<i>TATAGGTTGT</i>	<i>ATATTTCCGC</i>	<i>CATGATAATG</i>
651	<i>AGAGTTTCCCT</i>	<i>CTTTGAGATG</i>	<i>TTATTAAATC</i>	<i>ATTTCTCTCT</i>	<i>CGATATGTGT</i>
701	CAGGTTAACC	AAATTGGCAG	<u>TGTGACCGAG</u>	<u>AGTATTGAAG</u>	<u>CTG</u>

## 7.5 Partial potato genomic sequence of gene *Glo*<sup>b</sup>

1	<i>CTGATTTTCT</i>	<i>CCATTCATGC</i>	<i>TGTTGTTCCA</i>	<i>GGTTGTGAAA</i>	<i>GGGCACAAGG</i>
51	<i>CAGGATCCCT</i>	<i>GTATTCTTGG</i>	<i>ATGAAGGTGT</i>	<i>CCGCCGTGGA</i>	<i>ACAGATGTCT</i>
101	<i>TTAAAGCTTT</i>	<i>GGCACTTGGA</i>	<i>GCTTCAGGCA</i>	<i>TTTTTGTAAAG</i>	<i>TACCATATAA</i>
151	<i>TGATTGGACC</i>	<i>TAGATTGGTG</i>	<i>GCACATAGCT</i>	<i>CAATCTTTTA</i>	<i>TCTTTATCTA</i>
201	<i>CTCACCAGTA</i>	<i>GACCAAGTAG</i>	<i>TGCCTTTTTTC</i>	<i>CAAAATCTGC</i>	<i>CTTGAGAATG</i>
251	<i>GTCAACAATC</i>	<i>TGAAAGTGAT</i>	<i>TTTAAATCTC</i>	<i>CCTAATAGTC</i>	<i>GTAGTAATAC</i>
301	<i>AATAATTACG</i>	<i>CCTCAGTCCA</i>	<i>AAGCAAGTTT</i>	<i>AGTATCTTCC</i>	<i>CTTTTAAGCC</i>
351	<i>GCTATATTGG</i>	<i>GAAATTCAAA</i>	<i>GTTAATAGAA</i>	<i>GTCCCTTGCA</i>	<i>CTTTTCGAGAA</i>
401	<i>CCAAGAAAAG</i>	<i>TAGGCATTAT</i>	<i>AATTTATACC</i>	<i>TACATAATCA</i>	<i>GTGAATGATA</i>
451	<i>AATAACAGCT</i>	<i>GGTAAAATAT</i>	<i>TGGAGAACAA</i>	<i>GTTCCATGGG</i>	<i>TAAATGGGAA</i>
501	<i>ATGGGACAGT</i>	<i>CTTAACACAT</i>	<i>TCCAAAACCA</i>	<i>GTTTCATTGTT</i>	<i>GGCTAAGCAC</i>
551	<i>ACAGTTACAT</i>	<i>ATTATGCCAT</i>	<i>GTTTCCTTTT</i>	<i>CGTCTTCAAA</i>	<i>CATCTTAGCG</i>
601	<i>AGTGACATTC</i>	<i>TGTATTGTCC</i>	<i>TTAACAGATT</i>	<i>GGAAGGCCAG</i>	<i>TAGTTTTCTC</i>
651	<i>ATTGGCTGCT</i>	<i>GAAGGAGAAG</i>	<i>CTGGAGTCAA</i>	<i>AAAAGTGCTG</i>	<i>CAAATGTTGC</i>
701	GCGATGAGTT	TGAGCTAACT	<u>ATGGCATTGA</u>	<u>GTGGCTGTC</u>	

## 7.6 Partial potato genomic sequence of gene *Pgk*<sup>a</sup> (see Figure 3-4 on page 59)

\* Italics residues indicate intron; primer sequences were underlined; 'N' represents potential bi-allelic SNP with alternative nucleotides given in parentheses.

<sup>a</sup> Sequence data based on PCR product as template

<sup>b</sup> Sequence data based on cloned PCR product as template

*Rca*= ribulose biphosphate carboxylase activase, *Ant*= adenylate transporter, *Eno*= enolase, *Ndpk*= nucleoside diphosphate kinase, *Glo*= glycolate oxidase, *Pgk*= phosphoglycerate kinase

## ERKLÄRUNG

Ich versichere, daß ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; daß diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; daß sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie daß ich eine solche Veröffentlichung vor Abschluß des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Priv.-Doz. Dr. Christiane Gebhardt (MPIZ, Köln) betreut worden.

Köln, den 24. 08. 2000

CHEN Xinwei



---

### Teile dieser Arbeit sind in folgenden Veröffentlichungen enthalten

Xinwei Chen, Ralf Schaefer-Pregl, Josef Strahwald, Francesco Salamini and Christiane Gebhardt (2000) Function map as an approach to identification of genes controlling tuber starch content in potato cultivars. Plant and Animal Genome VIII (PAGVIII), January 9-12, 2000, San Diego, USA

Xinwei Chen, Francesco Salamini and Christiane Gebhardt (2000) A potato molecular function map for carbohydrate metabolism and transport. Theor. Appl. Genet. (in press)

## ACKNOWLEDGEMENT

I would like to thank Prof. F. Salamini who opened the Ph.D. position for me to work at the Max-Planck-Institut.

Many thanks and gratitude to Dr. Christiane Gebhardt for her excellent, enthusiastic and imaginative supervision. Her insight, advice and encouragement was most helpful throughout the three years. Special thanks to Dr. Ralf Schäfer-Pregl who patiently conducted the statistical analyses and answered all questions regarding statistics and to Dr. J. Strahwald who provided the phenotypic data.

I note particularly the help of Dr. Agim Ballvora, Dr. Ralf Schäfer-Pregl and Dr. Jonathan Phillips for their critical reading of the thesis, Juergen Paal for his excellent translation of the summary into German and Dr. Katharina Schneider for her helpful advice and encouragement.

I'm also indebted to Birgit Walkemeier, Heike Henselewski, Alexandra Niwergal and Werner Lehmann for their technical assistance.

Thanks are due to all of the group fellows for their considerable support throughout my study.

Finally I express my deepest appreciation to my wife who endured the long evenings and weekends and provided constant encouragement and to my son who brought us enjoyment and liveliness to the family.



**MPIZ-Köln**





**Name** CHEN Xinwei  
Kolibriweg 14  
50829 Köln

**Geburtsdatum und -ort** 07.09.1964 Tianmen, Hubei Province, China

**Nationalität** Chinesisch

**Eltern** Vater CHEN Xiaohong  
Mutter HE Wang

**Familienstand** verheiratet

**Schule** 1971-1976 Primary school  
Middle school  
High school

1977-1981 Institute of Genetics  
and Agronomy, China  
Plant Genetics,  
: Plant Physiology, Cell biology

1982-1985 Research Associate  
of Vegetables and Flowers  
Academy of Agricultural Sciences  
, China

1986-1988 Research Associate  
Avignon, France

1989-1991 Plant Genetics  
Research Associate  
DLO, Wageningen, the Netherlands

1992-1995 Plant Genetics  
natürlich-Naturwissenschaftlichen  
ität der Universität zu Köln  
anck-Institut für Züchtungsforschung  
Deutschland

